REPUBLIC OF TURKEY
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
DEPARTMENT OF EDUCATIONAL SCIENCES
CURRICULUM AND INSTRUCTION PROGRAM

DOCTORAL DISSERTATION

THE IMPACT OF FLIPPED CLASSROOM MODEL ON EFL LEARNERS' ACADEMIC ACHIEVEMENT, ATTITUDES AND SELF- EFFICACY BELIEFS: A MIXED METHOD STUDY

ORHAN İYİTOĞLU
13706004

ADVISOR
Prof. Dr. YAVUZ ERİŞEN

&

CO-ADVISOR
Prof. Dr. AHMET SABAN

İSTANBUL
2018

REPUBLIC OF TURKEY
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
DEPARTMENT OF EDUCATIONAL SCIENCES
CURRICULUM AND INSTRUCTION PROGRAM

DOCTORAL DISSERTATION

THE IMPACT OF FLIPPED CLASSROOM MODEL ON EFL LEARNERS' ACADEMIC ACHIEVEMENT, ATTITUDES AND SELF- EFFICACY BELIEFS: A MIXED METHOD STUDY

ORHAN İYİTOĞLU
13706004

ADVISOR
Prof. Dr. YAVUZ ERİŞEN

&

CO-ADVISOR
Prof. Dr. AHMET SABAN

İSTANBUL
2018

REPUBLIC OF TURKEY
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
DEPARTMENT OF EDUCATIONAL SCIENCES
CURRICULUM AND INSTRUCTION PROGRAM

DOCTORAL DISSERTATION

THE IMPACT OF FLIPPED CLASSROOM MODEL ON EFL LEARNERS' ACADEMIC ACHIEVEMENT, ATTITUDES AND SELF-EFFICACY BELIEFS: A MIXED METHOD STUDY
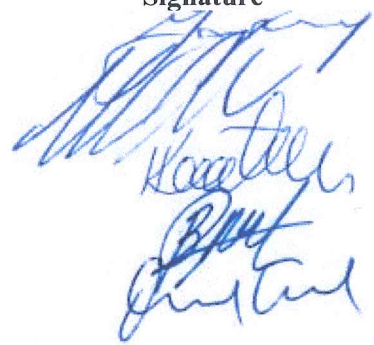
ORHAN İYİTOĞLU
13706004

Thesis Submission Date : .... /.... / 2018
Thesis Defense Date   : 25 /.01 / 2018
Approved unanimously / by a majority of votes

	Appellation, Full Name	**Signature**
Advisor	: Prof. Dr. Yavuz Erişen	
TMC Member	: Assoc. Prof. Mehmet ŞAHİN	
TMC Member	: Assoc. Prof. Hakan KARATAS	
Committee Members	: Assoc. Prof. Banu YÜCEL TOY	
Committee Members	: Asst. Prof. M. Onur CESUR	

İSTANBUL
JANUARY, 2018

ABSTRACT

THE IMPACT OF FLIPPED CLASSROOM MODEL ON EFL LEARNERS' ACADEMIC ACHIEVEMENT, ATTITUDES AND SELF-EFFICACY BELIEFS: A MIXED METHOD STUDY

Orhan İyitoğlu

January, 2018

The unprecedented pace of technology has inevitably manifested itself on practices and organization of the educational processes shaping not only the bridges of interaction that students build between themselves and their teachers but also the way they learn. As a response to the advent of new technology and the needs of these digital native students who grow up communicating through different means of technology, flipped classroom model has emerged as a form of blended learning that exploits the benefits of technology, constructivism, collaborative, distance learning and differentiated instruction by reversing the sequence of in and out of classroom procedures.

Grounded in a mixed method embedded design, the study has taken a distinctive step to explore the quality and efficiency of flipped classroom model in enhancing university prep students' overall academic performance in EFL and that in its sub-skills, retention of that performance, attitudes toward and self-efficacy beliefs in EFL.

Quantitative data was gathered through the administration of EFL Achievement Test, Attitudes toward EFL and Self-Efficacy Beliefs in EFL scales to 41 EFL students enrolled at Foreign Language School, Gebze Technical University in two different classrooms randomly assigned as experimental (N= 21) and control group (N=20). The intervention lasted during the whole 2016-2017 fall term. On the other hand, qualitative data was collected through follow-up semi-controlled interviews with 9 experiment group students from different achievement groups.

All the quantitative data was analyzed using SPSS 21 for Windows, LISREL 8.54 for Windows and ITEMAN4 while qualitative data was analyzed manually by employing content analysis procedures.

Findings of the study revealed flipped classroom as a teaching model significantly more effective than traditional lecture based instruction in enhancing EFL performance, long-term retention of this performance and improving its two vigorous predictors, attitudes toward and self-efficacy beliefs in EFL in higher education in Turkey. This facilitating role of flipped classroom model was also verified by qualitative results.

To conclude, delving into the important dynamics of EFL performance in Turkey, the present study has promised a bulk of valuable results that set flipping EFL classrooms as an efficient way of dealing with failure in EFL in Turkey.

Key Words: Academic Achievement in EFL, Attitudes, Blended Learning, Flipped Classroom Model, Self-Efficacy Beliefs.

ÖZ

TERS YÜZ SINIF MODELİNİN İNGİLİZCEYİ YABANCI DİL OLARAK ÖĞRENEN ÖĞRENCİLERİN AKADEMİK BAŞARILARI, TUTUMLARI VE ÖZ YETERLİK İNANÇLARI ÜZERİNDEKİ ETKİSİ: BİR KARMA YÖNTEM ÇALIŞMASI

Orhan İyitoğlu

Ocak, 2018

Teknolojinin baş döndürücü bir hızla gelişmesi sadece öğrencilerin birbirleri ve öğretmenleri ile iletişim kurma yöntemlerini değil ayrıca öğrenme biçimlerini de şekillendirerek eğitim uygulamalarında ve düzenlenmesinde kaçınılmaz olarak kendisini göstermektedir. Teknolojinin yeni boyutlarının ortaya çıkması ve teknolojinin farklı kanallarıyla iletişim kurarak büyüyen dijital açıdan yerli olan öğrencilerin ihtiyaçlarına karşılık olarak, ters yüz sınıf yöntemi teknolojiden, yapılandırmacılıktan, işbirlikçi öğrenme, uzaktan eğitim ve farklılaştırılmış öğretim yöntem ve yaklaşımlarından yararlanan, sınıf içi ve dışı prosedürlerini tersine çeviren bir çeşit harmanlanmış öğrenme modelidir.

Karma gömülü desene dayanan bu çalışma ters yüz sınıf yönteminin hazırlık öğrencilerinin yabancı dil olarak İngilizce öğrenmede genel performanslarıyla birlikte alt boyutlarındaki başarılarını, bu performansın kalıcılığını artırmadaki ve İngilizce öğrenmeye karşı tutum ve öz yeterlik inançlarını geliştirmedeki etkililiğini ve kalitesini ortaya koymayı amaçlamıştır.

Çalışmanın nicel verileri İngilizce Başarı Testi, İngilizce Dersine Yönelik Tutum ve İngilizce ile İlgili Özyeterlik İnancı Ölçekleri'nin Gebze Teknik Üniversitesi Yabancı Dil Okulu'nda okuyan 41 öğrenciye uygulanması ile toplamıştır. Bu öğrenciler çalışmanın değişkenleri açısından eşit olan iki sınıf rastgele deney (N= 21) ve kontrol grubu (N=20) olarak seçilmiştir. Bu deneysel uygulama 2016-2017 güz dönemi boyunca devam etmiştir. Öte yandan, çalışamanın nitel veriler deney grubunda farklı performans gösteren 9 öğrenci le yapılan yarı yapılandırılmış görüşmelerden elde edilmiştir.

Toplanan tüm nicel veriler SPSS 21, LISREL 8.54 ve ITEMAN4 programları kullanılarak nitel veriler de el ile içerik analizi yöntemi kullanılarak analiz edilmiştir.

Çalışmanın sonuçları ters yüz sınıf yönteminin düz anlatıma dayalı öğretim yapılan geleneksel sınıftan Türkiye 'de yüksek öğretim kurumlarında İngilizce başarısını artırmada, bu başarının kalıcılığını sağlamada ve bu performansın iki güçlü belirleyicisi olan tutum ve özyeterliliği geliştirmede daha etkili olduğunu ortaya koymuştur.Ters yüz sınıf yönetminin bu anlamlı etkisi nitel sonuçlarla da desteklenmiştir.

Sonuç olarak, Türkiye'de yabancı dil olarak İngilizce öğrenme dinamiklerinin derinine inen bu çalışmada, ters yüz sınıf yöntemini Türkiye'de İngilizce öğrenme

konusundaki başarızlıkla mücadele etmede etkili bir yol olarak dikkate alınması gereken sonuçlar paylaşılmıştır.

Anahtar Kelimeler: Yabancı Dil Olarak İngilizce Öğrenme Başarısı, Tutum, Harmanlanmış Öğrenme, Ters Yüz Sınıf Yöntemi, Özyeterlilik İnançları.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ANOVA   **:** Analysis of Variance

ANCOVA   **:** Analysis of Covariance

CFA   **:** Confirmatory Factor Analyses

EFL   **:** English as a Foreign Language

ESL   **:** English as a Second Language

FATIH   **:** The Opportunities and Improvement of Technology Project

IRB   **:** Institutional Review Board

ITEMAN   **:** Item and Test Analysis Program

LMS   **:** Learning Management System

MANOVA   **:** Multivariate Analysis of Variance

MOODLE   **:** Modular Object-Oriented Dynamic Learning Environment

PPT   **:** PowerPoint Presentation

SEM   **:** Structural Equation Modeling

SPSS   **:** Statistical Package for the Social Sciences

STEM   **:** Science, Technology, Engineering and Medicine

1. INTRODUCTION

The current study was designed in a mixed method embedded design to explore the efficiency of flipped classroom teaching model as solution of failure in EFL in Turkey. By dwelling on such vigorous predictors of EFL achievement in Turkish setting as attitudes toward, self-efficacy beliefs in EFL, the study was conducted to produce promising results to a wide range of concerned bodies. To realize this purpose, the first chapter was devoted to construct the basics of the study by embracing the dynamics that set failure in EFL as a significant problem, research problems and hypotheses shaped by the purpose of the study in addition to the assumptions based on the limitations. The chapter was finalized by drawing the conceptual framework of the study and defining common terms.

1.1. Statement of the Problem

In today's rapidly globalizing post-modern world, there have been a large number of changes and developments in every fields of the life ranging from science, technology to society, culture and more specifically human perceptions of one another. This process brought an inevitable impact on the educational processes with itself. As a result, learners began to be regarded as individuals as a whole construct of physical, cognitive and affective variables were carried into the heart of educational processes and procedures (Akbari, Hosseini, 2008). This new tendency has left its mark on the practices and perceptions of education itself. Within this perspective, the factors related to individual differences began to take up the most of the literature on teaching/learning and the professional literature was obsessed with terms and phrases which try to capture the elusive concepts that distinguish one person from another (Crozier, 1997).

In such an age of individualizing education, English as a Foreign Language Teaching (EFL), having important functions and roles, cannot resist those changes and developments triggered by the recent tendencies in education. One of the most important reasons behind this fact is that today's conditions have incapacitated those

who are able to communicate in only their native language in different aspects of their lives. Different reasons may prompt such an obligation to be competent in a second language. Yet, out of them, it seems that the necessity to adapt to the pace of the rapidly changing world and get integrated to the rest of the world attract most of the attention. In this multilingual and multicultural world, this communication bridge is constructed with an international language, English. This is mostly because, with the assumption of globalization process in the world, English is accepted as the most widely used foreign language (Crystal, 1997). Scientific, social, economic and technological developments accompanied by the use of English, adopting it as the official language by some organizations such as European Union and international relationships in various dimensions among the countries have also added to the crucial role of the language since the second half of 21st century (Guilherme, 2007). In this globalised world where it is accepted that English is the international language, using that language communicatively turns out to be first condition of being literate. Taking these crucial multilingual and multicultural aspects of EFL into account, a great deal of effort has been internationally performed to explore the sub-skills and the processes involved as the determinants of the success in learning it (Baker, Boonkit, 2004; Block, 1986; Brantmeirer, 2006; Clément, Dörnyei, Noels, 1994; Haley, 2004; Hosenfeld, 1997; Iyitoglu, Aydin, 2015; Javier, 1997; Keshavarz, Ashtarian, 2008; Kim, Wang, Ahn, Bong, 2015;Khajavy, Ghonsooly, Fatemi, 2017; Marefat, 2003; Oxford, Burry-Stock, 1995).

In addition to the aforementioned international studies, there have also been a number of studies conducted in Turkey focusing on the factors influencing the success and failure in EFL learning and teaching (Akdoğan, 2010; Akkuş, 2009; Aküzel, 2006; Ayhan, 1999; Çatal, 2015; Gömleksiz, 1993; Gülmez, 1982; Kabaharnup, 2010; Oğuz, 1999). These studies were carried out with different stakeholders ranging from teachers, students to administrators, parents from different cities and regions of the country. They all explore similar points that play a significant role in the failure of EFL learning in Turkey. These factors can be embraced under three categories as (Şahin, 2009):

a) Factors related to learning and teaching atmosphere
b) Factors related to students' perceptions about themselves and EFL
c) Factors related to family

When the results indicated by Şahin (2009) taken into account, it is understood that most of the reasons behind failure of EFL in Turkey is explained with the factors related to learning & teaching atmosphere and those related to students' perceptions about themselves as EFL learners. More specifically, while they explained 80, 2 % of those reasons behind unsuccessful EFL in secondary schools, they stood for 71, 2 % of the factors leading to failure in EFL in primary schools. To make the situation more visible, the participant teachers of the studies stressed the inadequacy of teaching materials, heavily reliance on traditional teaching methods (Kabaharnup, 2010). The results have also indicated that the participant teachers related the failure in EFL to mostly overcrowded classrooms and lack of modern technologies (Akkuş, 2009; Akyel, 2003; Özen, 1979). They also clarify the problem by linking the overcrowded classrooms to lack of necessary teaching time allocated to each of the students. This detection as one of the reasons behind most of the failure in EFL is significant since "all the learners need enough time to practice all the learned points to direct the brain to solve the codes of the new language" (Engin, Seven, 2007, 12).

As the means of opening the gates to the rest of the world, EFL loads learners with lots of responsibilities. However, these responsibilities may cause adverse effect on students' performance. In a similar way, a number of studies carried out on the reasons behind unsuccessful EFL learning and teaching in Turkey have stressed students' lack of interest, attitudes, motivation and self-efficacy beliefs in their EFL performance (Akkuş, 2009; Aküzel, 2006; Başaran, Cabaroğlu, 2015; Çatal, 2015; Kabaharnup, 2010; Kanadlı, Bağçeci, 2015; Kazazoğlu, 2013; Kiziltepe, 2000; Özen, 1979). As solutions to these problems, in those studies, the participant students and teachers of these studies suggested the integration of technology to enhance their motivation and belief in their abilities to be successful EFL learners.

Within this perspective, when the crucial roles and functions EFL carries in today's technology-driven globalizing world are taken into account, it seems as a must to take the steps to apply new teaching procedures. Those procedures need to be supplied with the technology and let the teachers increase the time they spend on teaching and tap into students' self-efficacy beliefs and motivation for learning English. Additionally, such a new model must be a context-specific remedy for the problems of EFL learning and teaching in Turkey. More specifically, the boundary of this model must be designed in accordance with the results of the aforementioned

studies conducted on the factors affecting the success of EFL learning and teaching in Turkey. Along with them, increasing tuition costs and fees, online and technology-embedded course offerings trigger the researchers to revise the wisdom of traditional teaching methods and to consider appropriate alternatives (Bishop, Verleger, 2013). The Flipped Classroom Teaching Model developed "out of a history of experimentation with the concept of hybrid, or blended learning and problem based learning, using active learning techniques and new technologies to engage students" stands as a response to the call of recent teaching model compatible with the needs of the age (Arnold-Garza, 2014, 8).

Also known as inverted classrooms, the flipped classroom-teaching model is simply defined as "events that have traditionally taken place inside the classroom now take place outside the classroom and vice versa" (Lage, Platt, Treglia, 2000, 32). In other words, the flipped classroom model implies re-design of in and out of classroom activities. While the students get the instruction at home through some formerly delivered means ranging from simple slides, podcasts, audio or narrated presentations to video casts including animations, screen captures, and other multimedia content, the classroom time is allocated to enhancing the related skills through problem-based teaching and engagement techniques (Educause, 2012).

The Flipped Classroom Teaching Model, cited mostly in the literature as based on student-centered theories, embraces interactive group learning activities inside the classroom, and direct computer-based online individual instruction outside the classroom (Bishop, Verleger, 2013). A number of international studies have indicated that the flipped model helps teachers in many ways. Thanks to it, teachers manage class time more efficiently by addressing multiple learning styles of the learners, donate them with more active learning opportunities getting them take responsibilities for their own learning, increase the time of interaction students and teachers spend on teaching and learning (Bergmann, Sams, 2012; Cole, Kritzer, 2009; Gallagher, 2009; Gannod, Berg, Helmick, 2008; Lage, Platt, Treglia, 2000; Overmyer, 2012; Siegle, 2014). As a result of these highlighted advantages of the application of the model in teaching, as stated by Arnold-Garza (2014), it is being applied in many disciplines ranging "from engineering to life sciences to business to statistics" (11).

Within this perspective, it seems that cited strengths of the flipped classroom model overlaps with the cited reasons behind the failure in EFL learning and teaching in Turkey. This study, therefore, incorporates the flipped classroom model and EFL teaching to answer the instructional needs and expectations of EFL learners and teachers in Turkey. In other words, by applying a flipped model rooted in blending distance learning with open-ended problem-solving experiences, the study takes the form of proposing a solution in response to main explored reasons of learners' being unsuccessful in EFL in Turkey. In this way, the present study attempts to shed lights on the impact of the flipped classroom model on EFL learners' academic performance, their self-efficacy beliefs in and attitudes toward EFL learning as compared to that of traditional teaching.

1.2. Significance of the Study

The unprecedented pace of technology has noticeably exerted its impact on many facets of today's global world ranging from commerce, policy, security to social and personal relationships. Over the last few years, this impact has manifested itself so strongly on education that the bridges of interaction the students build between themselves and their peers, and their teachers have changed prominently. As the hints of this change, many students today seem not to be able to lead their lives without their technological tools such as their mobile phones, I-pads, tablets and of course internet. Therefore, they bring most of them to their classroom. They communicate through different means by using social media sites like twitter, instagram and facebook. Moreover, together with reliance on such tools and means of social media, their ability to make use of technology are evolving rapidly (Johnson, 2013) and they turn into "digital natives" growing up with technology (Banitt, Theis, Leeuwe, 2013). In relation to this, all over the world, as a result of some studies, teenagers were found to spend an average of time ranging from 1-2 to 7,5 hours per day engaging in different types of media activities such as watching TV, playing video games, surfing the web, checking the news, using mobile devices, and listening to music (Li et al., 2007; Otrar, Argın, 2014; Rideout, Foehr, Roberts, 2010; Toruk, 2008).

Being aware of these changing tendencies of the learners as the members of the new huge global society, it is easy to see the fact that this new digitally surrounded

environment has also changed the way students learn today. Therefore, it has turned into a must to revise approaches to all parts of educational processes from teaching to learning in a way to exploit the benefits of technology, which seems to be a must for the digital natives of the age. Correspondingly, many countries around the world have taken serious steps to integrate technology into educational processes, too. As noted by Herold (2016), there has been a boom in education related technology and this industry exploits a slice of more than $8 billion market for hardware and software. For the education system to keep its pace with these changes, they simply replaced blackboards with smart boards and donated the learners with tablets and I-pads. Additionally, they utilized internet to be in touch with the students as much as possible through distance learning. In this international competition of installing the technological infrastructure of educational environments, Turkey has initiated with the third and fourth Five-Year Development Plans (Planning Commission, 1973; 1979). In those plans, the use of radio and television were stressed in non-formal mass education and it was decided to set up a TV channel for open distance university and mass education. In 1990s, some political documents were produced on science and technology. Since then, technology and internet have been more injected to educational processes in Turkey (Akıncı, Kurtoğlu, Seferoğlu, 2012). With the advent of The Opportunities and Improvement of Technology Project (FATIH), the country has taken more decisive steps for the integration of the technology into the education. However, teachers seem unable to keep pace with these changes. Despite a large number of technological opportunities, a number of research indicate that teachers seem not ready or willing to change the way they teach in both national and international context (Akkoyunlu, 2001; Altun, Ilgaz, 2016; Herold, 2016; Gilakjani, Sabouri, Zabihniaemran, 2015; Gürer, Tekinarslan, Yavuzalp, 2016). Therefore, there comes out the need for evidence based on research results that convince the teachers of the impact of technology on their students' learning outcomes.

In today's globally breathing world, it is clear that most of the life is driven by the technology related tools. According to the uncodified laws of the world, it is also certain that this technological globe is run through English. The use of English spreads at such a rate that it gets easier to witness its use in various places and fields such as international airports, sports, international mailing and conferences, commerce, international aid. According to Crystal (1997), there are approximately,

except for EFL students trying to learn the language, 1,2 billion non-native English speakers while the number of the people speaking it as their native language is 330 million. Stressing the number of the people, native or non-native scattered all around the world, Smith (2015) defines it as the language of the industrial revolution, science and technology. Supporting this claim, the results of the Research Trends (2012) study, which was conducted on 21,000 articles from 239 different countries indexed in SCOPUS, indicated that all non-English journals required the researchers to write their abstracts in English and % 80 of all were written completely in English. Much more surprisingly, the study also explored that in such countries as Germany, France, and Spain, papers written in English are more than those in the country's own language. As an example, in the Netherlands, this is 40 to 1.

In addition to the pervading share of English in science and technology, it seems that it is also common in the internet. According to the statistics of Internet World Stats (2016), English is the most used language in the internet. A study by W3Techs found out (2016) more than half of the most visited web sites have English homepages. When the reality that the communication among non-native speakers is generally held in English in international setting is taken into account, it will be better understood that being able to communicate in English equals with being literate today. Therefore, in addition to its official status in more than 70 countries and being taught as EFL in more than 100 countries (Crystal, 1997); English is clearly accepted to be much more global language than ever. In turn, it becomes more associated with students' academic performance in a number of different fields than ever. As a result, this makes it really crucial for the members of this global society to keep updated in English. That current situation makes English as a means of survival for teenagers who are explored to spend most of their time on their technological tools and the internet.

Despite the crucial role and functions English has in teenagers' lives, a body of research tragically point out the failure in EFL in Turkey (Akdoğan, 2010; Akkuş, 2009; Aküzel, 2006; Akyel, 2003; Ayhan, 1999; Başaran, Cabaroğlu, 2015; Çatal, 2015; Gömleksiz, 1993; Gülmez, 1982; Kabaharnup, 2010; Kanadlı, Bağçeci, 2015; Kazazoğlu, 2013; Kiziltepe, 2000; Oğuz, 1999; Özen, 1979). These studies conducted with different stakeholders in different settings attempted to shed light on the reasons behind the unsuccessful EFL in Turkey. While some of the reasons were

attributed to learning and teaching oriented atmosphere including the teachers, some of them were found to be related to those associated with the learners themselves. Out of them, commonly specified factors were found to be lack of positive attitudes toward and self-efficacy beliefs in EFL on the parts of the learners. On the parts of the teachers, they were specified as overcrowded classrooms lack of intelligent ways to integrate technology into teaching.

Within this respect, the Flipped Classroom Model comes out as a rescue to cope with the failure in EFL. Flipped Classroom Model, defined as one of the most important recent growth area in education, is believed to stand as a solution to most of the cited reasons behind the failure in EFL in Turkey. Flipping in and out of classroom procedures, this model promises to help teachers both tackle with the disadvantage of overcrowded classrooms more efficiently by allowing them to spend more time with their students and expand their students' skills through more problem-solving group work and task based learning activities in the classrooms. Therefore, this study investigating the impact of flipped classroom model on EFL learners' performance will produce valuable outcomes for a large number of people ranging from teachers, parents to curriculum developers and policy makers.

Understanding the role of media, internet and technology in teenagers' lives it becomes more important to enlighten those concerned about them including parents, the varying members of the world of education, policymakers (Rideout, Foehr, Roberts, 2010) Therefore, to deal more effectively with their problems, it is important to notice the digital gap between students' in and out of school lives (Downes, Bishop, 2012). This gap is formed because of teenagers' life routines. While they spend most of their times on the internet, playing computer games, listening to music, engaging with mobile phones or social networking out of classroom, they are exposed to traditional boring lessons lack of technology (Banitt, Theis, Leeuwe, 2013). Correspondingly, the flipped classroom model offers students online instructional videos out of the school and allocates class time for the expansion of the skills through problem-solving tasks. Right on this point, this model promises to tap into learners' interests by drawing on the sources they spend most of their time on. Within this respect, this study aims to shed lights on the impact of flipped classroom on learners' attitudes toward and self-efficacy beliefs in EFL. Therefore, by focusing on the cited predictors of success in EFL in Turkey, this study

offers significant results that must be taken into account by a wide range of concerned bodies.

Despite the flow of technology that will suggest a classroom model compatible with the needs of the 21st century, very little research is said to be conducted on flipped classroom (Abeysekera, Dawson, 2015). Since the flipped classroom stands for a recent topic in educational research, there seems to be lack of qualified research on its educational effectiveness (Bishop, Verleger, 2013). Supporting this, Abeysekera, Dawson (2015) claim to come up with, in contrast to its Google popularity, only eight articles in ERIC database including the phrase of the model in their titles, abstracts or keywords. They (2015) also stated only two of them were peer reviewed. In March, 2016, a similar search was carried out by the researcher of the current study himself in *Web of Science* database. On the database, 354 documents were found to include flipped classroom or inverted classroom in their titles. However, 119 of them were in the format of article written in English. Refining the research areas to educational sciences, the researcher explored only 82 of these articles written in English included "flipped classroom" or "inverted classroom" in their titles. The rest of those articles were in the areas ranging from chemistry, nursing and computer science to government law, surgery and agriculture. Moreover, the publication dates of the articles ranged from 2012 to 2016, which indicates flipped classroom as a recent research area of interest. In addition to students' attitudes toward flipped implementation, the studies were found to cover its effectiveness in a lot of different academic disciplines. However, as a result of search of the words "foreign" and "English" within the results, only 7 of them, published in 2015 to 2016, focused on flipped classroom model in English classrooms.

The rare effort to study flipped classroom model in the context of Turkey is also indicated by the researcher. In March, 2016, a similar search was undergone in Turkish Journal Park Academic and only 4 articles were explored to be published between 2015 and 2016. Yet, two of these were found to focus on students' opinions about the flipped classroom implementation while one of them was in the form of a theoretical review of the model and the other was a case study. In the same way, the search for the phrase in the database of Council of Higher Education thesis center found out only 11 thesis carried out on flipped classroom by March, 2016 (Akgün, 2015; Aydın, 2016; Balıkçı, 2015; Boyraz, 2014; Ekmekçi, 2014; Gençer, 2015;

Kara, 2016; Sırakaya, 2015; Turan, 2015; Yavuz, 2016; Yiğit, 2014). While three of them were master thesis, the other three were submitted for the Philosophy of Doctorate (PhD) degree. However, to the interest of the present study, only two of them were carried out in EFL. While one of the PhD study aimed to search the impact of the model on the success in writing in EFL (Ekmekçi, 2014), the other study was a master thesis designed to explore its impact on learners' performance in only two grammar structures (Boyraz, 2014). However, education is a long process (Ertürk, 1972) and as recommended by Boyraz (2014) the impact of an intervention must be traced in that long run. Therefore, the current study is significant since it aims to foster the understanding of this recent model on EFL learners' academic performance, attitudes and self-efficacy beliefs in the whole process of 2016-2017 fall term.

Thanks to its commonly cited benefit for enhancing lecture delivery, it is explored that the model is largely studied in disciplines such as science, technology, engineering, and mathematics (Hung, 2015). On the other hand, flip classroom model is theoretically claimed to be applicable in any subject area with students from different levels (Bergmann, Sams, 2012). However, it is currently applied with K-12 education (Horn, 2013). Therefore, there seems a need in the related literature to gain insights about the impact of flipped classroom on non-STEM higher education settings (Abeysekera, Dawson, 2015; Hung, 2015). Hence, this study will take a pioneering step to combine flipped classroom model and EFL in higher education in Turkey.

In brief, the present study embraces two important points in Turkey: Failure in EFL and the flipped classroom model. Drawing on key functions of English in today's world and young people's heavily reliance on technology and internet, the study aims to form a base from which concerned bodies ranging from parents, teachers to curriculum developers and policy makers examine the impact of the model on the determinants of the success in EFL in Turkey.

1.3. Purpose of the Study

The rapid globalization process has led to many changes in the science and technology engaging the countries into a never-ending competition of being a knowledge-based society (Erişen, 2007). Turkey, like many other countries, makes

efforts to take its own place in this competition incorporating different strategies and technologies into different parts of the life. On the way of being a knowledge-based society, with Increasing the Opportunities and Improvement of Technology Project (FATIH), it has started to set the basis on which internet, smart boards and some technological means such as tablets can be used in educational processes (Akgün, Yılmaz, Seferoğlu, 2011). This recent tendency will help the members of the society reach, use and process the information efficiently independent of the time and the setting (Sarıtaş, Üner, 2013).When the fact that the language of that rapidly changing information is English is accepted, bringing technology and the language together will also help the members of the society updated. Bearing this in mind, the present study will aim to provide insights about the impact of the flipped classroom teaching model on EFL learners' academic achievement and its retention, their attitudes toward and self-efficacy beliefs in EFL as compared to that of traditional teaching methods and strategies. In addition to eliciting students' personal opinions about flipped classroom teaching model, the possible impacts of it on their self-efficacy beliefs in and attitudes toward EFL will be further investigated through interviews with the students.

1.4. Research Questions

The aim to compare the impact of the Flipped Classroom and Traditional Teaching on EFL learners' academic achievement and its retention, their attitudes toward and self-efficacy beliefs in EFL is embodied in the following research questions:

1. Is there a statistically significant difference between pre- and post-test scores of the students in the traditional lecture-based classroom with regard to

   a) EFL Achievement Test,
   b) Sub-sections of EFL Achievement test
   c) Attitudes toward EFL Scale,
   d) Self-efficacy Beliefs in EFL Scale,
   e) Sub-scales of Self-efficacy Beliefs in EFL?

2. Is there a statistically significant difference between pre- and post-test scores of the students in the flipped classroom with regard to

   a) EFL Achievement Test,

     b) Sub-sections of EFL Achievement test

     c) Attitudes toward EFL Scale,

     d) Self-efficacy Beliefs in EFL Scale,

     e) Sub-scales of Self-efficacy Beliefs in EFL?

3. Is there a statistically significant difference between post-test scores of the students in the flipped classroom and traditional lecture-based classroom with regard to

     a) EFL Achievement Test,

     b) Sub-sections of EFL Achievement test

     c) Attitudes toward EFL Scale,

     d) Self-efficacy Beliefs in EFL Scale,

     e) Sub-scales of Self-efficacy Beliefs in EFL?

4. Is there a statistically significant difference in the retention of EFL achievement test performance between the students in the flipped classroom and traditional lecture-based classroom?

5. What are the EFL learners' perceptions of their learning experiences in the flipped classroom?

6. What are the EFL learners' perceptions of the impact of their learning experiences in the flipped classroom on their

     a) attitudes toward EFL and

     b) self-efficacy beliefs in EFL?

1.5. Research Hypotheses

With the aim of gaining insights into the impact of flipped classroom model on EFL, the following research hypotheses were listed to guide study:

1. The flipped classroom teaching model influences EFL learners' academic performance.

2. The flipped classroom teaching model influences learners' attitudes toward EFL.

3. The flipped classroom teaching model influences learners' academic self-efficacy beliefs in EFL.

1.6. Limitations of the Study

Every step was taken carefully by the researcher to carry out the study and protect participant anonymity and privacy. In this way, the findings of the study promise insights about typical Flipped Classroom implementations in different EFL settings and situations. However, the validity and reliability of the results are still exposed to some limitations. Within this respect, the followings can be listed;

1. This study is limited to 41 prep EFL learners in two different classrooms at Foreign Languages Department, Gebze Technical University in Kocaeli. Therefore, the number of the students in control and experimental groups made it impossible to generalize the results since a larger number of participants could have yielded to different and more accurate results.

2. This study is also limited to data collected through Attitude toward EFL Scale, Self-Efficacy Beliefs in EFL Scale, EFL Achievement Test administered and interviews held during the fall term of 2016-2017 academic year.

3. The study is limited to English classes.

4. Regarding the time and cost issues, a study conducted with students exerts some constraints out of the control of the researcher. The place where the surveys, interviews and achievement tests were employed, the time and duration of answering the questions and the type of assistance provided were some of the conditions which were beyond the control of the researcher.

1.7. Assumptions

In this study, it is assumed that;

1. The participant students in the pilot study will mark the best options reflecting their real EFL performance.

2. The participant students in the pilot study will mark the best options reflecting their real attitudes toward EFL.

3. The participant students in the pilot study will mark the best options reflecting their self-efficacy beliefs in EFL.

4. The participant students in the study will mark the best options reflecting their real EFL performance.

5. The participant students in the study will mark the best options reflecting their real attitudes toward EFL.

6. The participant students in the study will mark the best options reflecting their self-efficacy beliefs in EFL.

7. The participant students in the control and the experiment groups will not interact with one another during the study.

8. The participant students will answer the questions carefully and truthfully during the interview.

9. There will be no significant difference between the students in control and experiment groups in terms of variables within the focus of the study.

1.8. Conceptual Framework of the Study

The current study takes the form of scrutinizing the impact of flipped classroom teaching model on the EFL students' academic achievement, their attitudes toward, their self-efficacy beliefs in and their self-reported perceptions about EFL and the model itself through a mixed method embedded design. In this respect, this study shaped by the limits of the flipped classroom model.

Flip classroom model, a relatively new term in the literature, has been prevailing in education recently (Bergmann, Sams, 2012). However, this does not make this model a totally original (Hung, 2015). In the related literature, some other terms such as flipped classroom (Bergmann, Sams, 2012), inverted classroom (Lage, Platt, Treglia, 2000), inverted learning (Davis, 2013), just-in-time teaching (Novak, 2011) have been used to stand for the typical characteristics of this model for over a decade. On the other hand, the limited amount of literature and on the Flipped Classroom Model makes it difficult to draw the clear framework of the model. Yet, the common point of those differently termed versions of the model signal a specific pedagogical type of blended learning where conventional classroom sessions accompanied by homework are replaced and flipped in a way supported by instructional videos and enriched with problem solving tasks (Bergmann, Sams, 2012).

Flipped classroom teaching model is rooted in blended learning and based on the theoretical pillars of active learning (Hung, 2015). Set by *American Society for Training* in 2003 as one of the top ten popular trends and greatest trend in today's higher education, blended learning is simply defined as combining online and face-to face instructional methods (Graham, 2006). On the other hand, Staker, Horn (2012) expanded this definition putting blended learning as "a formal education program with face-to-face instruction, in which a student learns at least in part through online delivery of content and instruction, with some element of student control over time, place, path and/or pace" (Staker, Horn, 2012, 13).Graham (2006) grouped this into three types of blended learning on the level of their technology integration as enabling blends, enhancing blends and transforming blends. Current implementations of the flipped classroom model fits into the category of enhancing blends since it aims to provide enriched traditional classroom teaching with the integration of technology. However, it is critical to see that the flipped classroom model is different than *Technology-Rich Instruction*, *Distance Education* or "*E-Learning*" since those three only focus on the delivery of the content through internet (Staker, Horn, 2012).In these models, the educator either presents the content through smart boards, projectors or tablets in the traditional classrooms or delivers the content to the learners through internet not allowing face-to-face interaction. However, the flipped classroom model primarily focuses on the tenants of active learning but not content delivery (Bergmann, Sams, 2012).However, the flipped classroom model is deeply rooted in the active, face-to-face problem and inquiry based learning experiences within the classroom following the instructional videos outside the classroom with the aim of activating student preparation and allocating much more time for active learning activities in the classroom.

The flow of instruction in the flipped classroom model, consisted of two steps, is shown below in comparison with that of traditional teaching.

Figure 1.1: Comparison of Traditional and Flipped Classrooms

Farah, M. 2014. The Impact of Using Flipped Classroom Instruction on the Writing Performance of Twelfth Grade Female Emirati Students in the Applied Technology High School (ATHS). PhD Thesis, Dubai: The British University in Dubai, 13.

As seen in the figure, the flipped model is different than technology assisted classroom teaching with its heavy reliance on active learning. Active learning stands for an instructional method that engages students in doing things and reflecting on what they are doing (Prince, 2004). In this way, it serves as a broad umbrella term including a wide range of learning activities, instructional strategies, teaching methods that aim to actively engage students in the learning process (Hung, 2015). Within this respect, the flipped classroom model overlaps with active learning in that it aims to activate students' preparation before class through videos or presentations and engage them in the classroom instruction through open ended problem solving activities.

EFL learning problem in Turkey is fundamentally regarded as an instructional problem based heavily on the quality of teaching process. However, this problem can be better understood and solved when also looked at it from the perspectives of the learners. This takes the researchers to study the efficiency of learning process itself (İlhan, Karataş, 2015). Correspondingly, in the literature, it will be reviewed that positive attitudes toward and self-efficacy beliefs in EFL are significantly related to EFL achievement in Turkey (Akkuş, 2009; Aküzel, 2006; Başaran, Cabaroğlu, 2015; Çatal, 2015; Kabaharnup, 2010; Kanadlı, Bağçeci, 2015; Kazazoğlu, 2013; Kiziltepe,

2000; Özen, 1979). Accordingly, Csizér and Dörnyei (2005) explored attitude as a significant factor in learning a foreign language in the study where they investigated. Similarly, Gardner (2006) claims that foreign language achievement is not only based on one's ability but also motivation to learn that language. Motivation includes the perception of the learners toward learning a language (Yurtseven, Alcı, Karataş, 2014). These show attitudes as important in the amount of academic effort exerted during the process of learning the target language. As a result, the study is theoretically based on Gardner and Lambert's (1972) claim that positive attitude toward learning a foreign language can enhance the language learning process exerting impact on the foreign language learners' behaviors toward not only the target language but also its culture, and the community.

In addition to attitude, self-efficacy is another affective factor playing an important role in EFL performance (Yanar, Bümen, 2012). Research carried out during the last twenty years proved self-efficacy as a strong predictor of academic achievement (Rahimi, Abedini, 2009). Drawn from Social Cognition Theory, Bandura (1997) simply defines self-efficacy as "an individual's judgment about his or her ability to accomplish a given task or activity" (1). Bandura (1997) supports this impact of self-efficacy on academic performance explaining that it determines amount of the effort one expends to achieve something and emotional reactions toward something. In this respect, the current study is also grounded in the self-efficacy beliefs drawn from Bandura's social cognition theory seeing it as a crucial variable in language learning performance.

Within this perspective, the current study is conceptually rooted in flipped classroom model grounded in blended and active learning, Gardner and Lambert's (1972) focus on the role of attitude and self-efficacy beliefs in social cognition theory in language performance. By accepting their key roles in EFL achievement, the current study combines flipped classroom model with attitudes toward and self-efficacy beliefs in EFL. In this respect, the current study is based on the premise that the flipped classroom-teaching model facilitates learners' positive attitudes toward learning the language and strengthens their beliefs in their ability to be successful language learners. According to this assertion of this study, despite a lack of agreed definition of the concept of flipped classroom, in this study the flipped classroom is adopted as

a set of pedagogical approaches put forward by Abeysekera, Dawson (2015, 3) as follows:

> (1) move most information-transmission teaching out of class
>
> (2) use class time for learning activities that are active and social and
>
> (3) require students to complete pre- and/or post-class activities to fully benefit from in-class work.

To sum up, this two folded framework of the study is illustrated as in the following:
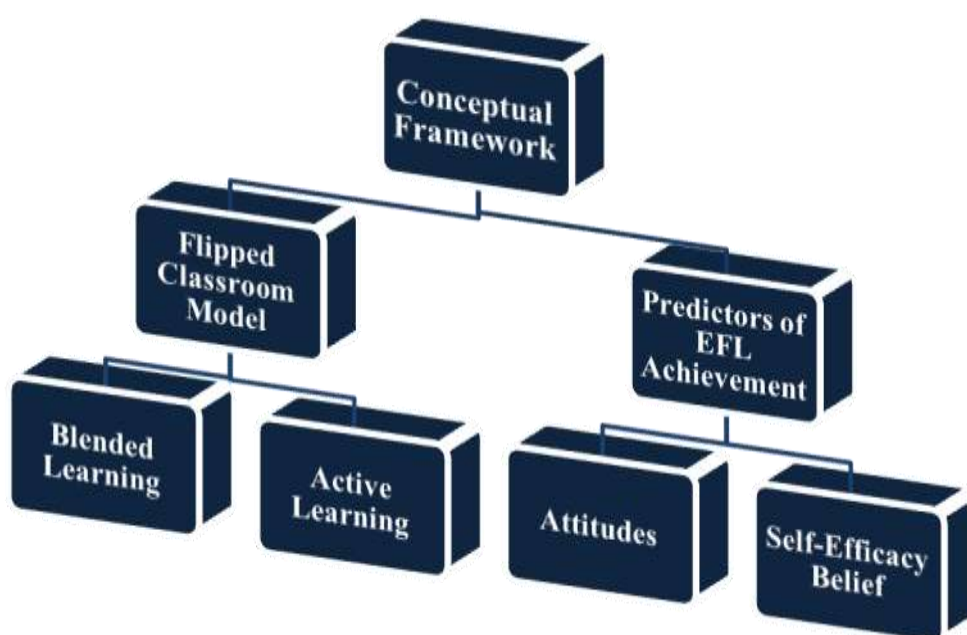
Figure 1.2: Conceptual Framework of Study

1.9. Definition of Terms

The major terms that are used frequently throughout the current study are defined as follows:

Flipped Classroom Model: A set of pedagogical approaches that allocates most direct instruction from the group learning space to the individual learning space by flipping the work at school with the work at home, as a result of which group space is transformed into a dynamic, active, interactive and social learning environment where the educator guides students as they apply concepts and engage creatively in the subject matter (Abeysekera, Dawson, 2015; Bergmann, Sams, 2012; Network, F. L., 2014;Lage, Platt, Treglia, 2000).

Traditional Teaching: A set of pedagogical approaches that is primarily based on direct, didactic or teacher-centered teaching during which the learners are frequently engaged with lower level learning activities in the classroom with the control of teacher. They are expected to elaborate on what they have learned in the classroom by doing "homework" at home after their lessons at school.

EFL Achievement: Level of attainment in EFL estimated by performance on the EFL Achievement Test.

EFL Achievement Test: A multiple choice test developed by the researcher employing the procedures of reliability and validity to measure university prep EFL learners' attainment of specified EFL skills (Duy, 2016)

Attitudes toward EFL: A concept that includes cognitive, affective and behavioral components referring to individual's beliefs, feelings and thoughts about outcomes or attributes of English as a foreign language and the act of learning it (Gardner, Lambert, 1972).

Attitudes toward English Scale: 5-point Likert-type scale including 19 items in four factors adapted from Tuncer, Berkant, Doğan (2015) rooted in the scales developed and adapted respectively by Aiken (1979) and Tunç (2003) in order to explore EFL learners' attitudes toward learning EFL.

Self-Efficacy Beliefs in EFL: Psychological construct that stands for beliefs in and personal determination of one's own capabilities to achieve a desired goal by organizing and executing specified courses of a certain task or activity related to learning EFL. (Bandura, 1997)

EFL Self-Efficacy Belief Scale: 5-point Likert-type scale adopted from Yanar, Bümen (2012). It includes 34 items in four factors, to explore EFL learners' self-reported beliefs in their knowledge and abilities to accomplish a task in four main ability domains of English as listening, reading, speaking and writing.

Semi-Structured Interview Form: A written form consisted of seven semi-controlled questions developed by the researcher to gain insights into the experiment group students' perceptions about their flipped classroom experience and its impact on their attitudes toward and self-efficacy beliefs in EFL.

2. REVIEW OF LITERATURE

This part of the study will touch on the basics of the study in two folds. Within this respect, the first section of this part will take the form of the attempt to provide insight into the basics of the flipped classroom model. Therefore, the researcher will shed lights on the literature related to Flipped Classroom covering its definition, foundations throughout history and its theoretical basics. Then after, the research studies conducted on flipped classroom practices in varying disciplines in national and international scope will be reviewed. The second fold will briefly explore attitudes and self-efficacy beliefs as the predictor of success in EFL in Turkey covering their theoretical boundaries and review of related literature on them both in international and national contexts. As a result of the review of literature, the necessary basis will be formed on which the present study will seek the impact of flipped classroom treatment on the predictors of academic achievement in EFL in Turkey.

2.1. Flipped Classroom Model

The second part of the review of literature will serve for empowering the structure of the study. In this respect, in addition to the previous parts on attitudes and self-efficacy in EFL, this part will delve into the flipped classroom model as a strongly significant trigger of predictors of EFL success. With this in mind, the research will detail the model covering its definition, historical development, theoretical framework, criticism directed to it with review of a number of studies on it from different disciplines all around the world.

2.1.1. Defining Flipped Classroom Model

Recently, "flipped" appears as a buzzword in today's student-centered education (Brame, 2013). However, there seems to be "a lack of consensus on what exactly the flipped classroom is, and there is also a limited amount of scholarly research on its effectiveness" (Bishop, Verleger, 2013, 5). Bishop and Verleger (2013) point out that flipped classroom is usually a focus on group-based interactive learning activities

inside the classroom referring to student-centered learning theories of Piaget and Vygotsky. Variation in these activities simply results in a number of definitions of the model by different experts with different but similar viewpoints. Therefore, this part will be the form of the attempt to reflect on some of those that draw the basic lines of the flipped classroom.

Flipped classroom model simply stands for a procedure where the instructor flips the work at school with the work at home. However, similar to this, there are a number of possible definitions of the model. Flipped Learning Network (FLN) defines flipped learning model as a pedagogical approach

> "in which direct instruction moves from the group learning space to the individual learning space, and the resulting group space is transformed into a dynamic, interactive learning environment where the educator guides students as they apply concepts and engage creatively in the subject matter" (2014, 1).

Another, probably the simplest definition of the flipped classroom, in coordination with *inverted,* is proposed by Lage, Platt, and Treglia (2000). "Inverting the classroom means that events that have traditionally taken place inside the classroom now take place outside the classroom and vice versa" (32). Despite its simplicity, Bishop and Verleger (2013) warn that such definition of the flipped classroom is not suitable to the nature of model the researchers and educators understand from its practice since it only touches on the flipping organization of in and out of classroom activities. To elaborate on it, it can be said that "a successfully flipped classroom involves more than just recording didactic content and sending it to students before a lesson: the time spent in class should be more important than the videos" (Basal, 2015, 29).

Baker, on the other hand, popularized the concept in the conferences between 1996 and 1998, and in 1998 opened the way for the model to be referred as "The Classroom Flip". Baker (2000) attempted to invert the transmission of rote learning content out of the class and use the time in the class for the students to work more on the application of the content. Within this perspective, Baker (2000) structured the approach around four verbs: clarify, expand, apply, and practice. Welcoming the changes in the structure of the model due to teachers' different personality types, Baker (2000) well drew the boundaries of the model. He (2000) elaborates on the model as a set of procedures where the goals are moving from sage to guide, opening up more time active learning in the classroom by reducing the time spent on the rote

lectures in the class with a focus on higher order skills, motivating students to take control and responsibility of their own learning by creating more opportunities for them to learn from their peers.

Mazur's "Peer Instruction" is another step in the development of the concept of flipped classroom (Brame, 2013). As a modified definition of flipped classroom, Mazur inverts passive learning outside of the classroom while carrying active learning inside the classroom (Crouch, Mazur, 2001). Within this respect, this traditional flipped instruction is seen as a set of procedures where students get involved with the content outside the classroom and the assignment to ensure their preparation for the inside class activities. Class time, accordingly, is shared for mini lectures or conceptual questions about the content under study.

Strayer, as another contributor to the development of the model, (2007) indicated the common point of most of the practices of flipped classroom as the goal of creating an active learning atmosphere. In this way, drawing on Piaget's theories on active learning which occurs not as a result of copying an idea but as a result of working on it, Strayer (2007, 45) defines the flipped classroom as a practice "usually motivated by a desire to learn through active participation in the classroom. What exactly is meant by active participation? Is not all learning active, whether from a book, a lecture, or small group activity?".

Most credited contributor in the process of popularizing the concept of flipped classroom, Bergmann and Sams (2012) focus on the inverted sequence of the activities and accordingly define the model as a learning and teaching atmosphere where what "is traditionally done in class is now done at home, and what is traditionally done as homework is now completed in class" (13). In addition to their definition, they (2012) also assert that various definitions of the flipped classroom gathered around allocating class time for more active learning and for more direct contact with the students in problem solving tasks in the classroom. In this way, as one of the milestones in the evolution of blended learning, the flipped classroom model serves for the attempts to incorporate the elements of technology and active learning into a curriculum (Hamdan et al., 2013).

As a composition of all these definitions, Bishop and Verleger (2013) put the concept of flipped classroom "as an educational technique that consists of two parts:

interactive group learning activities inside the classroom, and direct computer-based individual instruction outside the classroom." (5). Excluding and rejecting the definitions that do not include administration of videos as an outside of the classroom activity, the researchers (2013) do not set the difference from those broad concepts of the flipped classroom suggesting extensive reading and implementing discussion sessions in class.

As a result of these definitions, it is highlighted by different educators and researchers that flipped classroom concept stand for 'flipping' or 'inverting' sequence of in the classroom and outside the classroom activities with a focus on active learning. Despite its deep relation with blended learning, blended classes implement online elements inside the classroom with teachers' direct contact with the students (Hamdan et al., 2013). To exclude many different forms of the concept including especially assignment procedures in order not to lead to any more complications or misunderstandings, in this study, following definition of the flipped classroom proposed by Abeysekera and Dawson (2015) is adopted. As a result, flipped classroom is seen in this study "as a set of pedagogical approaches that:

1. move most information-transmission teaching out of class
2. use class time for learning activities that are active and social and
3. require students to complete pre- and/or post-class activities to fully benefit from in-class work." (3).

2.1.2. The History of Flipped Classroom Model

The Flipped Classroom is a teaching model that has attracted a lot of attention in the world of education over the last years (Johnson, 2013). Although it is put forward as a new innovative teaching model, there are some who claim it is rooted in the applications of the late 1990s (Baker, 2000). Since then, a number of terms such as *just-in-time teaching, inverted classroom, inverted learning* and *flipped classroom*, all of which are based on blended learning, e-learning and active learning stressing the importance of learners' preparation before class (Hung, 2015). Drawing on these bases, Hamdan et al. (2013) state the development of the flipped classroom model as a cumulative process formed as a result of some educators and researchers such as such as King (1993), Mazur (1997), Anderson and Walvoord (1998), Lage, Platt, and

Treglia (2000), Baker (2000), Tenneson and McGlasson (2006), Strayer (2007), Khan (2012), and Bergmann and Sams (2012).

The idea "Flipping classes" based on a number of student-centered teaching models and approaches is put forward to date back to King's article in 1993 (Baker, 2000). In the article entitled as "From Sage on the Stage to Guide on the Side", King (1993) described the tendency of active learning common in 1990s and as result opened the doors for the flipped class approach (Bergmann, Sams, 2012). King (1993), stressing the pillars of active learning with a constructivist viewpoint, attracted the attention to the changing role of the educator as either a facilitator or a guide but not a simple lecturer.

Mazur is a pioneer educationalist researcher on "Peer Instruction", which is seen as one of the first instructional foundations of flipped classroom model (Hamdan et al., 2013). As a professor at Harvard University, Mazur drew the boundaries of the elements of during-class interaction. These were developed further to shape the in-class sessions. Then, Mazur (2009) revised his instruction by recording his lectures through videos making students able to prepare for the class in advance. As a result of this change in the way of instruction, Mazur (2009) explored this type of technology use allows students to perform better during peer instructions and gives the teacher more time to work with the students and on the higher order skills. As a result, he came up with a Peer Instruction model including learners' first exposure to the content before the class, ensuring their preparation by some assignments and allocation of class time to the development of higher order skills.

Flipped classroom has been applied in similar forms in different disciplines for many years (Brame, 2013). Barbara Walvoord and Virginia Johnson Anderson are exemplified by Brame (2013) as the promoter of this approach thanks to their proposal of the model in which students first learn before the class and focus on higher order skills in class. To ensure that the students did necessary preparation for the class, they were said to propose an assignment-model. As a result, they claimed the students would receive more productive feedback decreasing the heavily reliance on the lecturer.

In 2000, Lage, Platt, and Treglia carried out an experiment entitled as "Inverting the classroom: A gateway to creating an inclusive learning environment", and led the

pathway to the emergence of flipped classroom with a similar approach as *inverted classroom.* Seeing that traditional classrooms do not appeal to every learning style the learners possess, the researchers (2000) designed an inverted classroom in an introductory economic course. With this in mind, the researchers donated the learners with the opportunity of many different ways of exposure to the content before the class such as textbook readings, videos, voiced PowerPoint presentations. Similar to previously stated applications, the researchers ensured students' preparation by some randomly graded worksheets. Class time was allocated to focus on the application of the learned content through problem solving tasks.

As another milestone, Baker (2000) took the step to reflect on the material that can be used in flipped classrooms. Motivated by King's (1993) work, Baker (2000) presented his paper at the 11th international conference on college teaching and learning in Florida, USA. In this presentation, Baker's (2000) attempted to enlighten the educators on adopting qualified programs and materials to invert, reverse or flip their classes.

Following these steps, Tenneson, McGlasson (2006) and Strayer (2007) have been the pioneers of drawing the boundaries of the model flipped classroom model fed by the results of a scientific study. In "*The classroom Flip*", Tenneson, McGlasson (2006) determined the limits of the structure of the flipped class approach in a curriculum. They (2006) explored that technological devices in teaching and learning process include a number of benefits for the students ranging from simply entertainment to networking. They further concluded that a teaching and learning environment enriched with media cultivates learning (Tenneson, McGlasson, 2006).

The earliest academic discussion of the 'flipped' classroom located is claimed to be Strayer's (2007) doctoral dissertation (Abeysekera, Dawson, 2015). Strayer (2007), on the other hand, in his dissertation titled as "The effects of the classroom flip on the learning environment: A comparison of learning activity in a traditional classroom and a flip classroom that used an intelligent tutoring system", which is one of the first research on flipped classroom, investigated how students perceived the learning environment and activity in the flipped and traditional classroom. Strayer's (2007) reached significant results that would shape the structure of the model. Accordingly, the researcher (2007) stressed the coordination of in and out of class activities for the model to be efficient and successful. Specifically, Strayer (2007)

explored that students in the flipped classroom "preferred and experienced a higher level of innovation and cooperation in their classroom" (106) while they were not content with the design of the class itself. Moreover, large number of activities in the flipped class "contributed to unsettledness among students (a feeling of being "lost") that students in the traditional classroom did not experience" (180). As a result, these two studies act as milestones in the development of the model since they offered results that must be taken into account while shaping the structure of the model in a way that would accept the voice of the learners.

The next step that makes the flipped classroom material common was taken by Khan and his academy. This story started in 2004 when Khan started to record math videos for his younger cousin having trouble at math (Khan, 2012). Since he did not have enough time to help his cousin, he created videos which he himself defined as a "virtual tutor" (Khan, 2012). Seeing that the videos helped his cousin take the control over her own learning, Khan founded the Khan Academy, and his collection of videos are chosen by some educators in flipped classrooms since they are usually well constructed and supported by clear explanations. As a result, this academy took another step to form the flipped classroom model entailing that professionally made educational videos can help students master their own learning.

Finally, two rural Colorado chemistry teachers, Bergmann and Sams are regarded as the pioneers of Flipped Learning in 2007 (Hamdan et al., 2013). These teachers started this journey of flipped classroom with a concern to compensate drop-outs and missed end-of-day classes of their secondary school students in their science classes due to the events they participated in. With this in mind, they recorded live videos and used special software to record their lectures or slides. To deliver the instructional content and videos of their class lectures to those students, they uploaded them on a YouTube channel to make the students able to access themes whenever and wherever they wanted. But they later explored that students who were not missing classes watched the videos as well in order to reinforce and revise the presented content. Seeing it, Bergmann and Sams (2012) have taught flipping their secondary school science classes since 2009. Based on their work with their students, they (2012) have published a book entitled as "*Flip Your Classroom: Reach Every Student in Every Class Every Day*". Although termed as inverted classroom or flipped instruction in the previous decades where assignments were used as a

preparation for in-class sessions, the flipped classroom is a term commonly and currently used in K-12 thanks to Bergmann and Sams (Farah, 2014). With its flexibility helping the educators design their instruction in a way to spend more time with their students on higher order skills in the class, the flipped classroom approach is still on its way of development as it is adopted and adapted by the educators all over the world (Bergmann and Sams, 2012)

2.1.3. Theoretical Framework

The flipped classroom model is fundamentally based on blended learning and active learning (Hung, 2015). Within this respect, it covers a large body of the active, face-to-face problem and inquiry based learning experiences within the classroom preceded by the instructional videos outside the classroom in order to activate student preparation by employing much more active learning activities with the students in the classroom. In this way, it is rooted in a type of instructional method that provides

> "an opportunity for students to use their new factual knowledge while they have access to immediate feedback from peers and the instructor, the flipped classroom helps students learn to correct misconceptions and organize their new knowledge such that it is more accessible for future use" (Brame, 2013, 3).

Engaging students in doing things and reflecting on what they are doing opens a pathway for the students to take control of their own learning. (Prince, 2004).

Standing for a kind of an "active blended" learning model where the learners deepen and facilitate the factual knowledge and its application, the theoretical framework of the flipped classroom model is clarified by the research team formed as a result of a cooperation between Flipped Learning Network and Pearson's School Achievement Services (Hamdan, et al., 2013).As a result of a review of literature on the flipped classroom, this cooperative research team identified the four key features underpinning the theoretical framework of the flipped classroom model approach in the form of pillars. Within this respect, the following part will be allocated for the four pillars of F-L-I-P^{TM}: Flexible Environment, Learning Culture, Intentional Content, and Professional Educators (Hamdan, et al., 2013, 5-6).

2.1.3.1. Flexible Environment

"F" of F-L-I-P^{TM} stands for flexible environment which is prerequisite for the implementation of the model since the "educators often physically rearrange their

learning space to accommodate the lesson or unit, which might involve group work, independent study, research, performance, and evaluation" (Hamdan et al., 2013, 4). In this way, the students are ensured that they can choose the time and place of their own learning taking more control of it. As opposed to traditional learning climate which is arranged in the form of the rows of the desks enriched with only boards and tables, the teachers of the flipped classrooms create flexible environments comprised of small individual and group work areas where most of the class time is allocated for the expansion of the skills through more problem based learning activities. This type of flexibility leads itself to flexibility in the teachers' "expectations of student timelines for learning and how students are assessed" (Hamdan, et al., 2013, 5).

2.1.3.2. Learning Culture

"L" of F-L-I-P^{TM} stands for a changing tendency in the learning concept from teacher centered approach to the student initiated one. Protecting the balance between motivation and de-motivation and covering readiness and level or zone of proximal development, Flipped Learning Network reflects this pillar as an instructional approach where students are not exposed to the teachers as the expert provider of the content but to the meaningful learning activities that help them explore the content much more deeply. In other words, the flipped classroom learning culture requires an inquiry-based teaching environment, where the face-to-face class time is processed not in a teacher-centered space but in a student-centered space (Bergmann, Sams, 2012). In this type of instruction, students are actively engaged with the knowledge construction process during which they "can theoretically pace their learning by reviewing content outside the group learning space, and teachers can maximize the use of face-to-face classroom interactions to check for and ensure student understanding and synthesis of the material" (Hamdan, et al., 2013, 5). As set by Brame (2013), the key in this learning is that "students are using class time to deepen their understanding and increase their skills at using their new knowledge" (3).

In brief, dominant student centered learning culture of flipped learning loads the teachers providing students with differentiated instruction facilitated by immediate feedback in the learning spaces shaped with the use of a great deal of media that leads to keeping a balance between direct instruction and scaffolding (Bergmann, Sams, 2012; Hamdan,et al., 2013).

2.1.3.3. Intentional Content

"I" of F-L-I-P^{TM} refers to the approach employed to organize the content itself in order to better address to students' learning needs. Flipped Learning Network determines the boundaries of this third pillar as the attempts to decide what content to be taught directly and how to organize it through instructional videos in a way that will allow students to "gain conceptual understanding, as well as procedural fluency" (Hamdan, et al., 2013, 6). In order to allocate most of the classroom time to helping students deepen their not only conceptual knowledge but also practical fluency of the content, instructional videos are employed as the means where the students are introduced to the content itself individually outside of the classroom. Therefore, to integrate the goals of the content with those of the learners, the content must be intentionally tailored to support the curriculum in a collaborative and active learning atmosphere to facilitate construction, application and synthesis of the content knowledge (Bergmann, Sams, 2012).

Having aforementioned critical functions into account, the instructional videos are formed by Flipped Learning Network (Hamdan et al., 2013). Accordingly, it was indicated by the cooperation between Flipped Learning Network and Pearson's School Achievement Services that the videos must be limited to the content that the students will use throughout the course necessary to finalize the projects. Moreover, assignments should be made into a video lecture; otherwise students may become confused and see the videos as a drawback to a course (Strayer, 2007). Students are also advised not to be confused with a large number of videos (Hamdan, et al., 2013).

In brief, the third pillar of the flipped classroom learning touches mostly on the attempts to prepare a meaningful basis of prior knowledge on which the students will blossom in a specific topic and tailor the whole instruction as to spend of the classroom time on promoting acquisition and application of the knowledge in addition to the incentives for them to prepare for the class.

2.1.3.4. Professional Educators

"P" of F-L-I-P^{TM} is the last pillar of the flipped classroom model standing for the teachers as the implementer of the approach as to ensure that students acquire and apply the learning content taking control and responsibility of their own learning. By donating the students with different opportunities to gain the conceptual knowledge

and produce intended learning outcomes at further levels of Bloom's taxonomy, "teachers shift direct learning out of the large group learning space and move it into the individual learning space, with the help of one of several technologies" (Hamdan, et al., 2013, 4). This corresponds to the role provided by the cooperation between Flipped Learning Network and Pearson's School Achievement Services for the teachers to take on in the flipped classrooms. Respectively, rather than respond deductively to the students as the source of the knowledge, adopting significant but less visible roles as professional educators adopt in the flipped classrooms, teachers facilitate the learning process and tailor the environment for the learners addressing their needs and preferences as much as possible. Within this perspective, professional educators of the flipped classrooms are advised by the cooperation to "continually observe their students, provide them with feedback relevant in the moment, and continuously assess their work by being reflective in their practice, connecting with each other to improve their trade, accept constructive criticism, and tolerate controlled classroom chaos" in addition to designing the units of the content, plan learning outcomes (Hamdan, et al., 2013, 6). Being able to proceed at their own pace in this way, students' zone of proximal development (ZPD) is not also threatened and they do not feel demoralized (Farah, 2014).

In brief, accepted key element of the flipped classroom practice signals professional guidance provided by the teachers for the students (Hamdan et al., 2013). This instructional foundation of the model is supported by Headden's (2013) assertion that, "the question is not really whether online instruction is a superior medium in and of itself. The Department of Education's survey found that success with online-learning depended on time spent on instruction, as well as the quality of the curriculum and pedagogy. In other words, it is not just the technology that counts but it is what educators do with it" (Headden, 2013, 17).

Last but not the least, in addition to mentioned pillars of the model referring to student-centered, collaborative, and interactive; the model is mostly based on the foundations of active learning, peer instruction, priming, pre-trainers and diversity of learners (Hamdan, et al., 2013). All these foundations and pillars combined with the expertise of teachers at drawing the framework of the content, creating, selecting and implementing appropriate videos and projects serve for pluralizing students'

opportunities to deepen their learning at further conceptual and practical levels by consciously moving direct instruction out of the classroom.

2.1.4. Criticism of Flipped Classroom Model

Education, with its significant functions effective in many facets of modern life, is a huge area where many individuals in and out of politics have focused their attention. This attention is embodied with a focus on school test scores, student achievement, and the demand for highly qualified teachers in the classroom (Douglas, Burton, Durham, 2008). Consequently, this dragged a great number of psychologists, researchers, academicians and educators to introduce new ideas into education. As one of the most recent steps taken to transform education, flipped classroom model has basically appeared as a result of the effort to inject technology into educational processes with an eclectic approach combining constructivist schema and behaviorist principles including active, problem-based learning activities and direct instructional lectures.

Gardner (1997) indicates that attempt to introduce new ideas to education brings some predictable consequences such as reaction and confusion. As with many other new born theories, flipped classroom model is also faced with some immediate criticism. However, it simply aimed to integrate technology into education in order to facilitate the acquisition of target skills and concepts before classroom session so that both students and teachers can benefit more from the public education. Reflecting on this purpose of the model, the first criticism is directed to its origin. In other words, although it is put forward as a new innovative teaching model appeared as result of Bergmann and Sams' efforts, two rural Colorado chemistry teachers, in 2007, November and Mull (2015) claim that the concept of flipped learning is not original but rooted in the Eric Mazur's and King's applications of the 1990s. Elaborating that some other educators have been using versions of flipped classroom model with minor differences since earlier than 90s, November and Mull (2015) manifested the reason behind this debate on the origin of the model as the lack of clear cut definition of flipped learning. This brings about an oversimplified definition that degrades the model to the combination of videos watched at home prior to class time and assignments completed in the classroom at school. However, such definition of the model focuses only on the organization of flipping the activities in and out of the

classroom (Bishop and Verleger, 2013). Therefore, it is useful to follow a definition of the model whose boundaries are well drawn around some concepts such as those set by Baker in the form of procedures (2000).

The flipped classroom is basically based on the efforts of the educators, teachers and researchers interested in such disciplines as science, technology, engineering, and mathematics (STEM) mostly in middle school environments (Arnold-Garza, 2014). Similarly, the majority of the related literature is focused on the impact of its application on exploring its use is focused on K-12 education (Horn, 2013). As a result of this tendency, another criticism is directed to flipped classroom for its limited scope of applicability and feasibility. However, Arnold-Garza (2014), in her study where she touches on the main characteristics of the model with examples from higher education, posits that there are a number of flipped classroom applications in higher education recently. This response to the speculated drawback of the model is supported by one of the pioneering and most cited studies ever conducted in the literature by Lage, Platt, Treglia (2000), with the aim of inverting their economics classroom. This criticism seems partly related to the misconception that the flipped classroom must cover the whole curriculum and the result of a static pedagogy. However, it is theoretically suitable to be tailored to match the specific needs of a wide range of learners in any subject area at different stages of education from primary to higher education (Abeysekera, Dawson, 2015; Bergmann, Sams, 2012; Hung, 2015).

Rooted in these two main areas of criticism, the review of the related literature explains challenges of the flipped classroom model in four categories: lack of access to required electronic devices and the internet, lack of time, lack of the skills on the parts of the teachers to apply the model and the students to be academically successful with it and the tendency to employ traditional teaching approaches.

2.1.4.1. Lack of Access

Flipped classroom model becomes prominent with its heavy dependence on the integration of technology into educational processes. However, one of the most common pitfalls of the model seems to be the problem of lack of access to technological tools required for the implementation of the model (Hanover Research, 2013). Nielsen (2012) who asserts that many students do not have access to such

technology at their homes and schools shares this disadvantage of the model. As a result, the model creates inequities among students for the disadvantage of those living in underprivileged conditions. On the other hand, this kind of argument may not be consistent with the realities of today's technology driven world since it becomes easier day by day for even the millions of people living in poverty to attain technology (Arnold-Garza, 2014).

Despite these problems of access to necessary technology due to socio-economic conditions, it is teachers' responsibility to list those who lack of necessary technological standards employing kind of survey before the implementation of the model (November, Mull, 2015). Accordingly, as a solution, teachers can multiply the delivery of the content. In other words, they can save the multimedia content onto CDs or DVDs, which can give the students lacking internet or even computers the opportunity to have an access to the content easily from home DVD players (Bergmann, Sams, 2012). As another solution including the former one, November and Mull (2015) suggest that by understanding their communities and working in coordination with both local libraries and community, schools can make the delivered content easily accessible for the students. As a result of this discussion about the lack of access to technology as an important pitfall of the model, it can be contended that such barriers can be easily solved through an elaborate planning made by flipped institutional bodies to create possible alternatives for those lacking access.

2.1.4.2. Lack of Time

Flipped classroom model depends on the use complex lecture tools embedded with the technology. This aspect of the model increases the load on teachers' shoulders and obliges them to spend a great deal of time to prepare the content for the flipped class (Arnold-Garza, 2014). This characteristic of the flipped classroom model is highlighted as a problematic issue (Bergmann, Sams, 2012; Lage, Platt, Treglia, 2000). In addition to spending time on preparing the technological lectures together with the exercises and the materials that will be used in the classroom, teachers also need to allocate much more time to learning skills to create required video content for the courses (Educause, 2012). Moreover, as set by Arnold-Garza (2014), planning of the whole course including the preparation of video lectures, assignments and adapting students to this new way of teaching and learning demands a substantial

amount of time. However, this aspect of the model indicated as a downside can be easily challenged. As a solution, it can simply be asserted that recycle and reuse of the content materials after the first implementation will reduce the time spent by the teachers on creating materials for the future flipped classrooms (Arnold-Garza, 2014). On the other hand, it is not true to expect that teachers must prepare the whole content for their classrooms. Regarding this point, November and Mull (2015) state that school leaders or administrators can hire those who are able to combine content knowledge and technology to create attractive videos. In addition to a group of professionals working to create the video content that will be employed by the teachers in the flipped classrooms, commercially purchased or free content prepared by a website cooperation such as Khan Academy, including a number of professionally made educational videos in different content areas, can be another way to challenge the specified disadvantage of the model.

2.1.4.3. Lack of Skills

Apart from the necessary skills to create the video content for the courses, the flipped classroom model requires the teachers to have and develop the ability to use the technology and the learners to have the literacy media skills to learn from the media (Educause, 2012)

In the first fold of the necessary skills for the adequate implementation of the model, the flipped teachers do not need to be able to create professional video content for the courses (November and Mull, 2015). However, they must possess the ability to choose from the ready-made videos and modify and adapt it to their learners (Bergmann, Sams, 2012). Moreover, teachers must be aware of some concerns if they decide to flip their classrooms. To be aware of such points ranging from the quality of the video recordings, their content consistency to their length, teachers must be made familiar with technology and educational media tools through a professional development program (Baker, 2000). As a result, to administer the basic point of the model adequately, teachers need to undergo a program that will donate them with the skills supporting not only their professional development but also their personal development.

In the second fold of the skills required by the model, the students are seen critical to have, develop and process the literacy media skills to reach, locate and use the media

(Hamdan, et. al., 2013). When the place and function technology possesses in today's life into account, it seems easier to accept that it is a must for the learners to have a medium level of media literacy skills in order to meet the demands of the globalizing world. Furthermore, the issue is much more critical for the flipped classroom model when its dependence on the technology is considered. As a result, it is suggested in the related literature that teachers must share time at the beginning of the term to explore their students' media skills and organize a plan to build and develop them to open the pathway for the success in the course (Arnold-Garza, 2014; Bergmann, Sams, 2012; Lage, et al., 2000). As a result, students will have a basis on which they can attain higher and more critical thinking skills that will help them synthesize new information and reflect on it in the real life problem solving tasks.

2.1.4.4. Challenge of Traditional Approach

The flipped classroom model is theoretically based on blended learning and active learning (Hung, 2015). Different from e-learning, it includes the implementation of active learning aided by face-to-face problem solving experiences during class time followed by instructional video content in the form of assignments outside the classroom. This concept of teaching fosters independent learning and allows students to reflect on their factual knowledge in meaningful problem solving activities (Brame, 2013; Prince, 2004). The related literature includes the examples citing the strengths of the flipped classroom model in different dimensions consistent with the goals it is set to pursue (Bergmann, Sams, 2012). However, these cited strengths may also turn into disadvantage for some other students. This disadvantage can be better seen in the probable mismatch between students' learning habits and flipped classroom model's media embedded teaching style. In other words, on the contrary to those who like to take more control of their own learning, some students may need more assistance since they may not be able to manage their own time and work (Arnold-Garza, 2014). Moreover, on the contrary to the fostering impact of pair work and meaningful activities, there are some situations when the students may want to work alone (Ash, 2012).

In addition to learning preferences of the students, the quality of the video lectures is also another point of discussion exerting an impact on the success of the learning and teaching process in flipped learning model. Firstly, as put forward by November and

Mull (2015), video lectures which last more than ten minutes and in which the students have no voice can passivate the learners in the flipped classroom. Due to a wide range of differences among learning styles of the learners, there may be some students who do not want to sit at their homes spending time on one-sided videos that aim to directly transfer the information. As a result, these videos, which do not allow students to interact with the information presented, may end up with a "bad pedagogy" where the teachers use them as the only means of instruction (Nielsen, 2012, 46). However, this dependence on the use of videos should not be processed by teachers to allocate more time to tests in the classroom. As a solution to this criticism, November and Mull (2015) suggest that creating interactive videos through online channels such as YouTube videos and Google forms that give a voice to the students prepared by different teachers in different styles will form a resource library tapping on different preferences over a period of time.

Similar to the learners, teachers tend to teach in the way they are already familiar with (Hamdan et al., 2013). Within this respect, teachers may tend to deliver the content to their learners not outside the classroom through online video lectures but in the classroom through face-to-face interaction. Together with an objection to the use of learning materials on some commercial education platforms, they may also administer videos not blended with the element of active learning (November, Mull, 2015). These tendencies may be overcome over a period of time as teachers witness the impact of flipped classroom model on their students' academic performances. Yet, their lack of ability to prepare professional videos can be challenged by either a kind of in-service training on efficient use of electronic tools or hiring a professional group of people able to create those videos (Arnold-Garza, 2014).

Stick deeply in traditional and direct teaching, teachers may feel less important than ever in the flipped classrooms since teaching there is realized outside the classroom through lecture videos (November, Mull, 2015). However, this is not the case since it requires a detailed planning on the part of teachers. Furthermore, as asserted by Nielsen (2012), the fact that flipped classroom model partly relies on a traditional model of teaching and learning does not mean that lecturing brings learning automatically with itself. Bergmann and Sams (2012) elaborate on this stating that videos cannot simply replace the teachers just for injecting technology into teaching. The success of it rather lies in its elaborate planning with blended approach where

pedagogical awareness directs the instructional processes without abusing the videos for delivering the content directly not to foster students' critical and independent learning skills (Miller, 2012)

As another criticism directed to flipped classroom model based on its concept of teaching, there seems to be the problem of accountability in the related literature (Arnold-Garza, 2014; Hamdan, et al., 2013; November, Mull, 2015). In other words, by heavily depending on their teaching and learning routines, teachers question how they will be certain that students have completed their assignments at home. This video lecture side of the model is also criticized since it steals students' time that they can spend socializing with their friends and families (Nielsen, 2012). However, this problem can be challenged by requiring students submit their reflections on each of the videos sent through answers to the questions or some small quizzes they take before they come to their schools (Johnson, 2013). Within this perspective, despite the tendency to follow the most familiar approach during learning and teaching process, teachers, over a period of time, may form their own library including a rich array of resources and donate the learners with efficient materials for revision and the opportunity to enlarge their skills since they will have more time to study with their mates and teachers in problem solving activities in the classroom (November, Mull, 2015).

In conclusion, despite some drawbacks of the model discussed above, the related literature also includes the examples that reflect the strengths of the model. They range from efficient use of class time to more active and independent learning opportunities for students helping them improve higher level thinking abilities (Bergmann, Sams, 2012; Cole, 2009; Gannod, Berg, Helmick, 2008; Lage, et al., 2000; Overmyer, 2012). In this respect, as stressed by Ash (2012), teachers must not degrade the model to that where students watch video lectures at home and do their homework more with the teacher in the classroom. Such a vision of teaching will not be able to change the way students used to learn revolutionarily. Therefore, a comprehensive awareness of the flipped classroom model rooted in blended and active learning with a constructivist approach adopted by the teachers is inevitable so that the model is best tailored for the needs of a specific group of learners (Nielsen, 2012).

2.1.5. Research on Flipped Classroom Model

Literature on flipped classroom model, also coded as inverted, asserts this recent technology-embedded-model as a specific type of blended learning design that is highly based on the administration of technology to help students learn the concepts of teaching outside the classroom and practise more on them in the classroom. With its characteristics that embrace interactive group learning activities inside the classroom, and direct computer-based online individual instruction outside the classroom employing active learning techniques and students-centered theories, flipped classroom model stands out as a remedy to be compatible with the needs of students in the 21st century. However, the significant place of the model in dealing effectively with technologically breathing students from all ages, especially teenagers, does not seem to be reflected in research carried out to locate its instructional effectiveness (Bishop, Verleger, 2013). Proving this lack of necessary qualified body of research in the flipped classroom model, a review of the phrase of the model, in the titles, abstracts or keywords of the articles in ERIC database by Abeysekera, Dawson (2015) indicated only eight articles despite its *Google* popularity. A similar search realized by the researcher of the current in *Web of Science* database limited to social and educational sciences, explored only 82 of 354 articles written in English covering "flipped" or "inverted" classroom in their titles. The publication dates of this limited number of articles ranging from 2012 to 2016 supported the claim that flipped classroom is a recent research area of interest. On the other hand, these research studies were explored to be largely rooted studied in STEM disciplines such as science, technology, engineering, and mathematics (Hung, 2015). Proving this claim, only 7 of 82 studies explored in *Web of Science* database by the researcher himself focused on flipped classroom model in EFL.

Within this respect, in order to locate the research on instructional effectiveness of the flipped classroom model in a wide range of disciplines including EFL, some studies carried out in international and national contexts were reviewed below. Accordingly, the first section of the following part was devoted to the review of the studies in the international scope starting with some pioneering and proceeding with more recent ones. In this first section, the studies conducted on the flipped classroom in EFL were preceded by the review of those conducted in other fields. Following

this, in the second section, the similar approach was adopted to review studies on flipped classroom model in Turkish context.

2.1.5.1. Research on Flipped Classroom Model in International Scope

Over the last years, there has been a great emphasis on creating learning and teaching atmosphere in various disciplines where students can spend more time on participating in more interactive activities that will tap on the development of high order skills (Kim et al, 2014). This type of atmosphere requires high level of peer collaboration on the part of the learners and rigorous planning on the part of the teachers to better respond to their needs and preferences in this digitally run period. Flipped classroom, in this respect, appears as a model rooted in blended and active learning that can meet these needs by flipping traditional classroom lectures with out-of classroom assignments to allocate more time on enhancing higher order skills (Strayer, 2012). Drawing on this framework of the model, there has been an increasing interest in the research on the flipped classroom model in a wide range of disciplines in education all over the world. Within this respect, this part of the study aimed to provide a review of literature on flipped classroom model in education from different aspects.

Attracting a high level of attention in the world of education recently, flipped classroom model is claimed to be rooted in the applications of the late 1990s (Baker, 2000). Baker (2000) is one of the first of those to carry out a research on flipped classroom model. Within this respect, Baker (2000) studied on the material that could be used in flipping his two courses, *Graphic Design for Interactive Multimedia* and *Communication in the Information Age.* With the use of an online content management, Baker (2000) get the learners reach lecture notes on a web page and work on the application of those contents in the classroom. In this study, Baker (2000) aimed to place the rote transmission of the content in the classroom with internet delivered presentations in order to provide the students take more control of their learning by spending more time on higher cognitive skills. His action plan based on four verbs as *clarify, expand, apply, and practice* was evaluated by students through responses. Participant students were asked to evaluate their flipped classroom experience with the other traditional courses they were taking. The results indicated positive perception toward their flipped classroom experience. Students'

comments also revealed that learning in two flipped classrooms was more individual helping them take more control of their own learning thanks to online materials and more active fostering cooperation with the peers through group activities in the classroom. Presenting the results of this study in a paper at the 11th international conference on college teaching and learning in Florida, Baker termed this method as *The Classroom Flip* (Johnson, 2013).

Following a similar procedure to that of Baker (2000), Lage, Platt, and Treglia (2000) carried out an experiment in their Introduction to Microeconomics course. Referring to the concept as *The Inverted Classroom*, the researchers (2000) asked the students taking the course to study on the videotaped content or PowerPoint presentations before they came to the school in order to have more time in the classroom to work on the problematic points with regard to the sent content and the application of that content in group activities. In order to provide compatibility between students' learning styles and the content, students were left free to read the textbook, watch a delivered traditional lecture, or viewing PowerPoint presentations or combinations of them. This study was carried out with 80 introductory economics students in five sections of the course. Starting with question and answer section for not more than 10 minutes, the courses were proceeded with an experiment, a lab, group activities in the classroom to respond to the diversity of learning styles among the students. A likert-scale followed by open-ended questions was used to locate students' perceptions toward their flipped classroom experience. The results evidenced positive reactions to the flipped classroom and its role in the increase in the peer collaboration and interaction. In addition to students' preference of flipped classroom to traditional classrooms, the results also revealed that flipped classroom, thanks to the presentation of the content in a number of ways, better matched with the diverse learner styles (Lage, Platt, and Treglia, 2000).

Two chemistry teachers, Bergmann and Samsare regarded as the pioneers and the founder of the recent term *Flipped Learning* (Hamdan et al., 2013). Though not a scientific research itself, Bergmann and Sams, in 2006, reported their observations about students' perception and achievement after adopting the model. These teachers opened the highway for the flipped classroom with the aim of dealing with the drop-outs and missed end-of-day classes of their secondary school students in their science classes. Therefore, they used podcasting and vodcasting their courses to help absent

students. They uploaded their instructional content on a YouTube channel to provide those students with an access to compensate their drop-outs. Later seeing that, students who regularly attended the classes also made use of these delivered content, increasing student interaction and peer collaboration more in the classroom, Bergmann and Sams stepped forward to flip their classes since 2009 and published books and articles on the model (Bergmann and Sams, 2012). Contributing to setting the term as Flipped Classroom in K-12, Bergmann and Sams were attributed as those who popularized the model rooted in the implementation of others preceding them. Since then, Sams and Bergmann, initiating the foundation of Flipped Learning Network™ (FLN), attempted to help educators providing them with necessary knowledge, skills, and resources to with regard to the successful implementation of model (Hamdan et al., 2013).

Strayer's (2007) doctoral dissertation is regarded one of the earliest academic discussion of the 'flipped' classroom. Strayer (2007) carried out his study to compare the flipped and traditional classroom based on the perceptions of students in two *Introduction to Statistics* classrooms at university. In this experimental study, one of those classrooms was designed according to flipped classroom model and therefore the content was delivered by an intelligent tutoring system outside the classroom. 23 students were engaged with each other to through active learning activities in the classroom. On the other hand, the other classroom was structured according to traditional lecture and homework design with the participation of 27 students. At the end of their learning experience, the students were asked to self-report their perception of their experience through *College and University Classroom Environment Inventory*. The data was enriched through field notes, mid-term and final exam notes, classroom transcripts, student interviews, student focus groups, researcher journal entries and student reflections. The results revealed that students in the flipped classroom reported better cooperation with their peers while they were not content with the design of the classroom, which was found to contribute to the unsettledness of the students. On the other hand, no impact of the model on students' success was explored. In this respect, this study by Strayer (2007) provided results that must be taken into account while shaping the design of the model and employing that model in the classroom.

The flipped classroom literature hosts a number of studies attempting to locate the impact of the model on students' academic performance in STEM in addition to their perceptions of the model (Arnold-Garza, 2014). In order to provide an example for those studies focusing on the discipline of engineering, the study conducted by Gannod, Burge and Helmick (2008) was reviewed. In this paper, the researchers adopted the term *inverted classroom* for the model. Within this study, the researchers proposed how to invert software engineering curriculum at Miami University. In addition to this, Gannod, Burge and Helmick (2008) reported the results of a piloted inverted course (*Service Oriented Architecture*) and their experiences of some other courses (*Fundamentals of Programming and Problem Solving and Data Abstractions and Data Structure)* that were being piloted at the time of the study in 2007. Podcasted lecture materials for the inverted course ranged from PowerPoint presentations and screen casts to video blogs. The results were acquired through the administration of *Small Group Instructional Diagnostic (SGID),* a technique generally used at Miami University to learn how students evaluate specific courses. The results yielded positive perceptions of the inverted classroom implementation. In specific terms, the students experiencing this piloting classroom appreciated the use of a number of techniques by the instructor and software demonstration podcasts as a valuable means of learning. As a result of this study, Gannod, Burge and Helmick (2008) asserted flipped classroom in engineering as a model exploiting the benefits of collaborative and distance learning.

Another large-scale experimental study conducted by Zappe et al. (2009) aimed to explore instructional effectiveness of the model in the discipline of architecture. In other words, this study was carried out in an undergraduate architectural engineering course with the participation of 95 students enrolled at *Introduction to the Building Industry* course. The researchers aimed to explore students' use of delivered video content and their perceptions of flipped classroom. Within the course, the teacher designed the instruction of one of three topics using iTunes video content in the spring term of 2008. Students were ensured to watch the videos through online quizzes based on the feedback gathered from the administration of piloted flipped classroom session before the study, the survey including checklists, rating items followed by open-ended questions was applied to the students. The results indicated that the students were eager to use their time out of the school to watch the delivered

instructional videos. Furthermore, the participant students also perceived flipped classroom had a positive impact on their understanding of the course content.

Another study that can be evaluated to meet flipped classroom model criteria was conducted by Moravec et al (2010). In this study, the researchers searched the impact of flipped classroom model on students' performance in a large introductory biology class for three of 30 period-classes at University of California in 2009.In this respect, the participant students were asked to cover *Learn Before Lecture* (LBL) material as termed by the researchers. These included watching narrated PowerPoint videos and completing a worksheet in advance to the course. This was done to use the time in the classroom for enhancing students' abilities to apply the content knowledge. The data was gathered through students' performance in the exam questions related to the flipped topic. The comparison of students' marks in the flipped content in 2009 with the traditional content in 2007 and 2008 explored that students' performance in LBL content was significantly higher. In more specific terms, this type of instruction resulted in the increase of 21 % in students' academic performance in biology.

Another study penned by Valenza (2012) was found valuable enough by the researcher of the current study to be shared in the review of related literature on flipped classroom. The reason behind this is the important figures included in the article. In the article entitled as "*The Flipping Librarian",* Valenza (2012) included the results of a survey administered by the Flipped Learning Network. The results of the survey evaluated the flipped classroom from the viewpoints of the teachers with regard to their job satisfaction, students' performance and attitudes. In this study, 453 instructors from different fields were surveyed. 88 % of them reported that flipped classroom fostered their level of job satisfaction while 67 % of them stressed that their students' performance at tests improved. Finally, 80 % of the participants expressed positive impact of the flipped classroom model on students' attitudes toward the courses and the learning content while the responses of 99 % of the educators indicated that they were eager to implement the model in their classes again next year.

Despite the density of research studies on STEM disciplines in the literature about flipped classroom model, there have been few efforts to conceptualize it in the context of medical education. In order to provide a review of related studies on a wide range of areas as much as possible, the current researcher handled the study

performed by McLaughlin et al. (2013). Within this respect, McLaughlin et al. (2013) aimed to search the impact of flipped classroom model on students' engagement, performance, and perceptions in a basic pharmaceutics course at two satellite campuses of the University of North Carolina. In this mixed method study, the researchers (2013) surveyed the participant 22 students for their perceptions and engagement behaviors at the beginning and end of the flipped course. The students were asked to watch the video lectures delivered to them through the course web site. For performance indicators, the researchers used the results of final exams from 2011 and 2012. This comparative study was conducted with the participation of the students taking the basic pharmacy course in traditional and flipped designs in 2011 and 2012 respectively. As a result, the researchers cited the enhancing impact of the flipped model over traditional model on students' learning. In other words, students' support for the learning content increased significantly stressing the positive impact of learning the content before coming to school on their ability to deepen their understanding of the related content. However, on the contrary to commonly cited result in the literature on the flipped classroom model, the researchers (2013) found no significant difference between students' academic performance in 2011 and 2012. Yet, qualitative data attained from the surveys revealed that the participant pharmacy students believed in the impact of the flipped model on their potentials to develop and be better engaged with the course.

Building on the gap in the related literature due to some limitations occurred by the design of flipped classroom itself, Kim et al. (2014) initiated a different comparative mixed method study with the participation of 3 instructors and 115 students enrolled in three separate classes in the departments of engineering, social studies and humanities at the University of Southern California in the fall term of 2012. In order to come up with some principles guiding a more comprehensive flipped design, the researchers aimed to explore students' and teachers' perceptions of the flipped classroom application. A four-likert scale with 50 items was employed to collect students' perceptions of the model. Then, students who were found to perform at two opposite poles were interviewed through a semi-structured protocol. Teachers' perceptions of the model, on the other hand, were gathered through the administration of *Instructors Reflection Protocol.* Based on the *Revised Community of Inquiry (RCOI),* the researchers analyzed the both quantitatively and qualitatively.

The results of the study showed students in general posited positive impact of the flipped model on their understanding of the content and their performance in the related area. However, students' positive perceptions of the model were significant than those in the departments of engineering and humanities. Centring the analysis of the collected data around such four sub-scales of RCOI *as* cognitive presence, social presence, teaching presence, and learner presence, the researchers proposed nine principles that guide the flipped model. As a result, Kim et al. (2014) suggested a broader concept model of the flipped classroom limited to exchange of in and out of classroom activities to help the teachers meet students' needs in a wide range of disciplines.

In STEM courses, there has recently been a popular increase in using flipped classroom model approach thanks to the advances in technology. The study carried out by Love et al. (2014) is an illustration of the reflection of this tendency into the dimension of mathematics. The researchers (2014), within this respect, aimed to compare traditional lecture model and flipped classroom model in applied linear algebra course with the participation of 55 students in two sections. The sections were randomly assigned as flipped and traditional lecture model implementation. Students' academic achievement was evaluated according to three common midterm exams and a comprehensive final exam while end-of-semester survey developed by the researcher was administered to measure their perceptions of the model. The results of students' performance in the exams revealed no significant difference between the students in the flipped classroom and traditional lecture model designs. However, students' performance in the flipped classroom had a more significant increase between the exams than those in the traditionally instructed classroom. On the other hand, the results of the survey indicated positive attitudes of the students toward flipped classroom implementation. More specifically, these students reported to enjoy peer collaboration aspect of the model stressing its positive impact on their ability to learn and remember better.

A study representative of common tendency to look for the instructional effectiveness of flipped classroom model or students' perceptions of it in the related literature was fulfilled by Gilboy, Heinerichs and Pazzaglia (2015). This research study is valuable since it incorporates flipped classroom model with Bloom's taxonomy. The main purpose of this study was to illustrate the basic framework with

regard to the implementation of the model and students' perceptions of it in comparison with traditional teaching models. Within this respect, 2 undergraduate nutrition courses were redesigned adopting the flipped classroom model in a manner to cover their three components termed as before class, during class and after class sessions. By a template provided by the researchers (2015), these 2 courses were ensured to include all levels of Bloom's taxonomy. While the first two steps were devoted to focus of lower levels of Bloom's taxonomy, the last component aimed to enhance higher order skills such as application, analysis, and synthesis. Four topics within 2 undergraduate nutrition courses were taught to 196 students at West Chester University following this design. 5-point Likert Scale developed by the faculty member having experience on the model was found to be reliable assessor of students' perceptions of flipped classroom model ($\alpha = .71$). 142 students taking the courses responded to the survey items. The results indicated students were highly satisfied with their flipped classroom experience. In addition to students' preference of instructional procedures followed during 3 sessions of the courses to traditional approaches, their responses to open-ended questions of the survey also reported that students liked to take responsibility of their own learning studying and proceeding at their own pace. By embracing the implementation of the model with a comprehensive approach covering Bloom's taxonomy, Gilboy, Heinerichs and Pazzaglia (2015) provide concerned bodies with important suggestions to ensure the success of the model.

Flipped classroom is mostly regarded in the literature as a teaching model compatible with the changing tendencies of today's students to learn and teachers to teach. In this respect, educational scholars continue to introduce its application in various disciplines. The study conducted by Long et al. (2016) is one of the representative studies in the recent literature on the implementation of flipped classroom model in educational context. This study takes the form of the effort to reflect on not only students' but also teachers' perceptions of and attitudes toward the use of the model in *The Technology-Enabled Active Learning* (TEAL) classroom. This classroom is defined by Long et al. (2016) as a "small capacity classroom equipped with multimedia projectors, white boards, laptops, and tablets, and utilizes modular tables for flexibly configured working arrangement" (46). An instructor and 55 students from different majors in science and engineering who took flipped TEAL course

participated in the study. The data was gathered from the instructor and 5 voluntary students from five different majors through the administration of a semi-structured interview. As a result of the thematic analysis qualitative data, the model was reported to assist students in increasing their learning motivation while enhancing their collaborative skills and problem-solving skills. On the other hand, despite students' and teacher's positive attitudes toward the flipped TEAL classroom, some critical suggestions were proposed by them about the design and organization of learning environment, technology. In other words, the participants' suggestions about the duration of the pre-class videos and warnings on the distractions that can be caused by the technological environment in the classroom may lead the flipped instructors to draw a framework that will enhance instructional effectiveness.

Review of recent literature on the implementation of flipped classroom in educational contexts has indicated that flipped classroom model has become popular at all educational levels with a significant density in STEM disciplines. The reason behind this may stem from heavy reliance of STEM disciplines on traditional lecture design to transmit knowledge (Hung, 2015). It seems that the promise of the flipped model to enhance lecture delivery by using classroom time more effectively through collaborative and problem-solving in-class activities to increase class preparation, interaction and performance attract the attention of scholars interested in STEM fields. However, scholars from social sciences or humanities attribute little importance to transmission of knowledge by trying to help their students construct their own knowledge since they already favour inductive teaching methods (Hung, 2015). Despite this, although limited, flipped classroom gains attention in some other fields ranging from medicine to social sciences (Reid, 2016). This insufficient interest is also reflected into EFL. Supporting this claim, the researcher of the present study explored in a search restricted to educational science research area in *Web of Science* that only 82 studies including "flipped classroom "or "inverted classroom" in their titles were articles written in English. A search for "foreign" or "English" within these results indicated only 7 of these studies were conducted about the flipped classroom regarding different aspects of EFL. Moreover, these 7 articles were found to be published in 2015 to 2016. In this vein, some of these articles were touched below to expand the review of literature on flipped classroom to non-STEM fields in a way that will include EFL.

One of the recent steps to introduce flipped classroom model into EFL was taken by Hung (2015). Hung (2015) designed a post-test-only quasi-experimental study in order to examine the impact of flipped classroom model on 75 EFL learners', taking a communicative course at a Taiwanese university in Taiwan, academic achievement, their attitudes toward the course and their engagement levels. In this respect, the students participated in three designs of flipped classroom termed as structured, semi-structured and non-flipped lessons. By the administration of a *Test of English as International Communication* (TOEIC), the participant students were found to have no significant differences with regard to their proficiency levels. Therefore, they were randomly assigned to three designs of classrooms. Each group took the lessons for three weeks. While *Web Quests* and *TED-Ed* tools were used for structured and semi-structured classes, traditional lecture model was followed with non-flipped lessons. The data for the study was gathered using end-of-lesson assessments given in the week following the lessons. Lesson study logs were employed after the completion of each lesson to let the students self- report time they spent on watching the videos in advance to the classrooms. On the other hand, a five-point likert learning experience questionnaire including 26 items and a semi-structured interview with 18 students from different achievement scores across three instructional groups were also employed. As a result of quantitative and qualitative analysis, it was explored the flipped classroom had a positive impact on EFL learning. In more specific terms, students in structured and semi-structured flipped classrooms performed academically better, developed more positive attitudes toward EFL while allocating more of their time to study before they come to school.

The experimentally designed study conducted by Leis, Cookie and Tohei (2015) took the form of another attempt to tap the issue of flipped classroom model in EFL. The researchers expanded the model to EFL in Japanese context and included 22 Japanese university students taking two English composition writing courses at two universities in Japan yielded results that indicate the positive impact of the model in EFL writing. As a pre-test, students were asked to watch a video in the classroom and write a composition about it. The number of the words they produced in this composition were recorded to compare with their post-test scores. Then, students were assigned to flipped and traditional writing classes following the same schedule and textbook. The experiment lasted ten weeks. During nine weeks, students in the

traditional classroom based on a keynote presentation lecture while slides converted into movie were sent to those in the flipped classroom on YouTube. This naturally differed the amount of time spent in the classroom on writing composition every week. At the end of this ten-week process, students were asked to repeat the same task they took as pre-test. Participants were also asked to complete a survey to have the data that shows duration of time they spent before and after the class each week during the course. To elaborate on the results produced by the study, the students taking the flipped writing course were found to spend significantly much more time on getting ready for the course. This difference was based by the researchers (2015) on the wide range of instructional options such as watching the video, practicing to write a composition or looking up a dictionary when compared to those in the traditional classroom whose options were limited to doing homework or flicking through the pages of a dictionary. Surprisingly, no difference between the time spent after the class was found although the students in the traditionally designed classrooms were expected to do their homework. This showed that students experiencing the flipped classrooms spent time after the class as much as those in traditional class. They were reported to watch the delivered videos again. In terms of the number of the words produced in the post-test, post-tests taken by students in the traditional and flipped classrooms were compared. This indicated students taking the flipped course wrote compositions including significantly higher number of words. Finally, comparison of pre and post-tests among control and experiment groups revealed significantly enhancing impact of flipped classroom on Japanese EFL students' writing skills.

A recent study performed by Hao (2016) contributed to the expansion of literature on flipped classroom. Approaching the issue from a different aspect, Hao (2016) focused on the concept of flipped classroom readiness covering the factors exerting an influence on it. In order to explore the relationship between flipped classroom readiness and a number of personal characteristics, 387 7th graders in an EFL classroom in Taiwan were included in the study. To obtain data that shows the participant students' readiness for the flipped model, a 5-point Likert *Flipped Learning Readiness Scale* (FLRS) including 27 items was developed by the researcher. The scale was found to base on five sub-scales (α = .75 to .92) as follows: "Learner control and self-directed learning, technology self-efficacy, motivation for

learning, in-class communication self-efficacy, and doing previews" (Hao, 2016, 297). The personal characteristics in the study were reported to cover gender, support outside the school and resources, use of internet, foreign language beliefs, and perceptions of English teachers. Within this respect, *the Beliefs about Language Learning Inventory (BALLI)* was used to measure students' beliefs about foreign language learning while the *Semantic Differential Scale* (SDS) was administered to find out their perceptions of their EFL teachers. The data with regard to other personal characteristics was gathered through some items added into the specified scales. SPPS was used by the researcher to conduct factor analyses, t-tests, one-way analyses of variance (ANOVA), the Pearson Product-Moment Correlation and multiple regression analyses. As a result of these analyses, it was explored that students were ready for the flipped classroom in a range from slightly above neutral to below-neutral. Compatible with the dynamics of this technological era, students' highest readiness was explored to be for technology self-efficacy sub-scale while the lowest was found to be for in-class communication self-efficacy and doing previews. Furthermore, sub-scales of BALLI were found to be significantly correlated with the sub-scales of FLRS. On the other hand, three sub-scales of SDS, teacher competency, teacher appearance and empathy, was found to be correlated with all the sub-scales of the readiness scale excluding technology self-efficacy sub-scale. In terms of other personal factors, support outside the school, availability of online learning resources, their language performance, study time and use of internet for fun were explored to predict high level of flipped classroom readiness while gender was found to have no impact on it. The results yielded a rich body of information to theoretically conceptualize flipped classroom in EFL and help EFL teachers design their classes in a way to embrace all the aspects of personal variables surrounding flipped classroom model.

Kvashnina and Martynko (2016) stepped further to analyze the potentials of flipped classroom in English as a Second Language (ESL) teaching setting. With this aim in mind, the researchers (2016) conducted an experimental study with students taking English for engineering course at Tomsk Polytechnic University in Russia. The analysis of the collected data yielded results compatible most of the related research studies. In other words, the study confirmed the positive impact of flipped classroom on students' performance, motivation and learning skills. To set the results in detail,

the students taking the flipped course scored on the tests 28 % higher than those taking the traditional course. In addition to the significant difference between their performances, the students in the flipped classroom got more motivated to learn with an enhancement on their learning skills. However, Kvashnina and Martynko (2016) warned based on the results that certain parts of ESL classes were suitable to be flipped, which leads the teachers to carefully analyze the course content.

To deepen the understanding about the implementation of flipped classroom in EFL, a qualitative case study conducted by Shaffer (2016). This study aimed to approach the issue of flipped classroom from the point of a teacher implementing the model. Within this respect, Shaffer (2016) conducted this research with an English teacher of more than 20-year experience and a master degree in educational technology using the model for the first time. This teacher was teaching 6 11th graders in U.S. literature class at a high school in the South western United States. To have a better understanding of the issue, Shaffer (2016) followed triangulation and collected data from interviews, field notes, and documents. The researcher had 3 semi-structured interviews with the teacher before, during and after the implementation. Documents were collected from the teacher's different materials utilized during the course. Observations of the planning of two-week unit on *The Great Gatsby* by F. Scott Fitzgerald and classroom interactions turned into data in the form of field-notes. Through the thematic analysis of the data, it was explored that many aspects of the teacher such as planning, use of technology, pedagogical posture and efficient management of time were affected by the flipped classroom model. On the other hand, the teacher in the study reported that the flipped model motivated the students fostering the classroom interaction.

Review of the literature covered above revealed that there is a buzz around the flipped classroom as a new teaching model in a wide range of academic disciplines at all levels ranging STEM fields and natural sciences to educational and the social sciences (Bishop, Verleger, 2013; Reid, 2016). Despite a few qualitative or mixed-method studies, most of the research followed an experimental design that assigned two groups of students to flipped and traditional classrooms. Participant students at different educational levels were engaged in tasks varying in type, length, and content in different fields. Despite some minor differences, the studies yielded consistent results with regard to the impact of the model on students' performance,

motivation, attitudes, perceptions, self-efficacy beliefs, learning skills and teachers' pedagogical views and their efficient use of classroom time. In this vein, the common thread in this review of literature is that flipped classroom model in different disciplines facilitates students' motivation to learn, attitudes toward their courses by enhancing their beliefs in their abilities to be successful in a learner-centered classroom atmosphere. Moreover, by helping the instructors create more classroom time to promote higher order skills by increasing the resources they have. However, flipped classroom is reported in the related literature not as a cure for all types of educational problems but as a strong potential to enable learning with a careful and an effective planning that will consider its possible constraints.

2.1.5.2. Research on Flipped Classroom Model in Turkish Context

Review of recent literature on flipped classroom in education has strongly indicated that the model is subject to significant amount of attention from a number of different fields. The review of this related literature explores significant contribution of the model to a wide range of educational environment (Uzunboylu, Karagozlu, 2015). However, these efforts, also clarified by the search on web of science for "flipped" or "inverted classroom" resulting in 82 articles from educational research area, are insufficient to "detail the design principles of the flipped classroom" and what aspects of flipped classroom implementations explicitly benefit teaching and learning" (Kim et al., 2014).

Similar to international literature, there seems to be a lack of studies on the implementation of flipped classroom model in Turkish context. This rare effort on flipped classroom model in the context of Turkey is proved by the researcher of the present study. The researcher, in this respect, searched for "flipped" or "inverted" classroom in the titles of the articles published in the journals indexed in *Turkish Journal Park Academic.* Accordingly, 4 articles (Basal, 2015; Kara, 2016; McKeown, 2016; Zeren, 2016) were explored to be published on the flipped classroom implementation in 2015 to 2016. Following a similar approach, the researcher enlarged the search to the database of *Council of Higher Education Thesis Centre*. As a result of this search, the researcher explored only eleven thesis were carried out in 2014 and 2016. Seven of these were master thesis while the other four were thesis for the Philosophy of Doctorate (PhD) degree (Akgün, 2015; Aydın,

2016; Balıkçı, 2015; Boyraz, 2014; Ekmekçi, 2014; Gençer, 2015; Kara, 2016; Sırakaya, 2015; Turan, 2015; Yavuz, 2016; Yiğit, 2014).At this point of the study, the researcher summarized some of these studies in order to provide a basis of studies in Turkish context on which the results yielded by the present study could be better evaluated.

In terms of the articles located in the database of *Turkish Journal Park Academic*, three articles were found to be published in 2016 while the remaining one was explored to be published in 2015. Of these studies, the one penned by Kara (2016b) was not a research study but a review of the related studies to cover all aspects of the model in education. Since this study produced no original data to be compared with the present study, it was not included among those studies conducted in Turkish context.

Another study conducted by Zeren (2016) took the form of the attempt to find out university students' perceptions of flipped classroom in a non-EFL field in Turkey. In this perspective, the study included 135 students from geography major at a university in Turkey. These participant students were taking four flipped courses in a geography programme in the Faculty of Science and Literature. The participants were provided with the resources electronically a week before the course. Accordingly, they were asked to watch PowerPoint presentations, videos, read the articles and answer the short exams by preparing extra question. In this descriptive study, the main source of data was gathered from students' compositions about flipped classroom model including their suggestions for improving its quality. This data was analyzed qualitatively adopting a thematic approach. As a result of it, 8 categories regarding the participants' observations on the advantages of flipped classroom and 4 categories regarding their suggestions for fostering its application were formed. Generally, the results indicated positive impact of the model on affective factors such as attitudes toward the lesson, anxiety and motivation. Furthermore, it was explored to facilitate students' cognitive abilities including critical thinking, planning and realizing inquiry, dealing with their misconceptions, independent learning and deepening understanding. On the other hand, the participant students' suggestion categorized in the study promised significant data that should be taken into account by the instructors or curriculum designer that aim to come up with an effective flipped classroom model. Accordingly, such a model is

reported to require to vary activities implemented in the classroom in order to address as many different learning styles as possible. It also requires to ensure all students have an access to the necessary technology and not to base all of the lecture on the presentations outside the classroom.

On the contrary to the aforementioned two studies (Kara, 2016; Zeren, 2016), the remaining two studies (Basal, 2015; McKeown, 2016) detected in the search for the related studies in Turkish context were explored to focus on flipped classroom in EFL. Compared to 7 EFL related flipped research of 82 studies in the international scope, the total number of studies seems dramatically low. Yet, the number of studies focusing on its application in EFL stands for 50 %. The first of these two studies was conducted by Basal (2015). In that study, 47 prospective EFL teachers taking *Flipped Advanced Reading and Writing I-II* course participated in 2012-2013 academic year. Similar to the study conducted by Zeren (2016), Basal (2015) aimed to explore participant prospective English language teachers' perceptions of flipped classroom model through a qualitative design based on content analysis of their responses to the question on the benefits of using video lecture in the specified course. Since students were given no information about flipped classroom, the question was worded as "What are the benefits of using video lecture in Advanced Reading and Writing II?" (Basal, 2015, 31). The implementation of the model lasted two terms at university in Istanbul. Based on the problems determined in the fall term, the model was redesigned. Accordingly, the students were asked to provide secret words included in the videos during in-classroom sessions to ensure that they watched the videos in advance to classroom. A learning management system (LMS) was used to deliver the lecture content, limited to 15-minute-videos, to the students at least four days before the lesson. Themes formed as a result of content analysis revealed positive impact of the model on EFL in different dimensions. First of all, it was analyzed to foster the quality of students' preparation for the course through the delivered professional content. Pushing the lecture content outside the classroom, the model was found to eliminate the constraints caused by the limitations of time. Allowing the students proceed in the course at their own pace by letting them watch the videos again whenever they want, the model is found to facilitate the interactive learning environment in the classroom. With these results, Basal (2015) thinks the success of the flipped classroom model lies in the teachers who implement and

stresses its potential to be used in different EFL courses with a comprehensive planning of outside and inside the classroom sessions.

Second and the last of studies reflecting flipped classroom model into EFL in Turkish setting was conducted by McKeown (2016). The researcher, in this study, adopted case study design to shed lights on the implementation of the model in the foundation year English Language Preparatory Program at a private university in Istanbul. In an environment where the learners were EFL learners and the teachers were native, the English Language Preparatory Program was redesigned by a team of 14 instructors. Since this team decided they spent most of their classroom time on grammar, they focused on creating resources for grammar presentations. As a result, they came up with 250 videos and the classes were equipped with a smart-board while each student was provided with a tablet. This new program was used with 419 students in 2014-2015 education year. Despite providing information on how to create resources and structure the assessment, the researcher (2016) did not mention about how the data was gathered or analyzed. Yet, the results were grouped into two as "successes and areas for development" (McKeown, 2016, 148). In terms of the successes of the model, the students were described to enjoy the use of technology as a natural part of their daily courses. Enabling to spend more of the classroom time on student-centered interactive tasks, the model was reported to exert a positive impact on students' collaboration and teachers' ability to manage the interaction in the classroom. Students' satisfaction was measured through online survey on whose reliability and validity procedures no information was shared. The results revealed students' satisfaction with the model ranging from 58 % to 95 %. 50 % of the students reported to study English outside the classroom ranging from 1 to 15 hours. In addition to its impact on the pass rate over the expectations at 84%", 70 % of the students reported that their English improved thanks to the model. The study elaborated also on the areas that should be considered to develop in order to help those interested in flipped classroom model in EFL setting. Within this respect, McKeown (2016) suggests that both students and teachers are provided with support to be effective users of technology through orientation sessions. This study, in this way, promises a rich body of information to redesign in-service teacher training.

In addition to four articles on EFL in Turkey, eleven studies, seven of whom were master thesis, (Akgün, 2015; Aydın, 2016; Balıkçı, 2015; Boyraz, 2014; Ekmekçi,

2014; Gençer, 2015; Kara, 2016a; Sırakaya, 2015; Turan, 2015; Yavuz, 2016; Yiğit, 2014) were found to be conducted on the implementation of flipped classroom model in different educational contexts in Turkey. The search for "flipped classroom" in the database of *Council of Higher Education Thesis Centre* revealed nine studies were conducted on non-EFL setting. Surprisingly, seven of these studies were explored to be conducted within the department of *Computer Education and Instructional Technology* at different universities in Turkey. The common thread in these studies was that Turkish scholars followed pre-test post-test quasi-experimental design in order to explore the impact of flipped classroom model on the participant students' academic performance, their learning motivation and their attitudes toward the attended flipped courses. In this part of the study, some of these seven studies (Akgün, 2015; Aydın, 2016; Balıkçı, 2015; Gençer, 2015; Sırakaya, 2015; Turan, 2015; Yavuz, 2016) were reviewed to represent the mentioned effort in the literature on flipped classroom in Turkey.

As an example of these seven studies centered around computer related programmes in Turkey, Balıkçı (2015) structured a pre-test post-test experimental design to determine the impact of flipped classroom on students' academic achievement in addition to their opinion about the flipped course. In this respect, 34 students taking *Web Editing* course at the *Department of Computer Programming* at the *Siverek Vocational School of Harran University* participated in the study. The course was instructed following a traditional and flipped model in control and experimental groups respectively during the spring term of 2013-2014 academic year. Compatible to the common findings in the related literature, *Independent-Samples t-test* showed significant difference between students' scores of achievement test. In addition to the impact of flipped classroom on the experiment group students' academic performance, the model was found to exert a fostering influence on students' opinions about their course.

Another study by Turan (2015) sought to reveal the impact of the flipped classroom implementation on students' academic performance, cognitive load and their learning motivation. In this comprehensive study devoted for philosophy of doctorate degree, Turan (2015) majoring in the department of Computer Education and Instructional Technologyheld a *concurrent triangulation design* to comprehensively detail the flipped classroom model in education. In this vein, following convenience sampling

method, 116 pre-service teachers, 58 in the control and 58 in the experimental group, from an Early Childhood Education department in Ataturk University were included in the study. The data for the study were collected through the administration of an achievement test, cognitive load and motivation scales, student view questionnaire and a semi-structured interview form. In the quantitative part of the study, Turan (2015) administered achievement test developed by the researcher as pre and post-test before and after the treatment. On the other hand, a 9 point Likert Cognitive Load Scale was used at the end of each course in the experimental group and at the end of each assignment in the control group. Furthermore, a 5 point Likert Motivation Scale was used as only post-test in both groups. In the qualitative part of the study to deepen the understanding about the flipped classroom model, Turan (2015) developed and administered a semi-structured interview form and student view questionnaire to the students experiencing the flipped model. The results of the analysis of the quantitative data revealed the positive impact of the model on students' scores of achievement test and motivational scale while exerting a negative impact on their scores of cognitive load scale. On the other hand, the analysis of qualitative data indicated that students' negative perceptions of the model when they were first told about it before the treatment were changed into positive after the treatment. As a result, fostering classroom interaction, retention skills, this comprehensive mixed method study supported the idea that flipped classroom model contributed to students' performance, motivation and perceptions over the traditional model rooted in lecture and homework procedure.

A thesis study for the philosophy of doctorate degree was performed by Sırakaya (2015). The purpose of this mixed method designed study was two folded. In this vein, Sırakaya (2015) first attempted to explore the impact of flipped classroom on students' academic achievement, self-directed learning readiness and motivation. In the second fold, students' perceptions of the model were aimed to be explored. This study was utilized with the participation of 66 pre-service teachers (34 control, 32 experimental) from Guidance and Psychological Counselling department at Ahi Evran University. These students were 4th grade students taking Scientific Research Methods course in the fall term of 2014-2015 academic year. The students in control and experimental groups were equalized based on the results of achievement test developed by the researcher. Additionally, academic achievement was measured

through quizzes scores, class activities scores. On the other hand, the data for self-directed learning readiness and motivation was measured by administrating a 5 point Likert Self-Directed Learning Readiness Scale and a 7 point Likert motivation scale as pre- and post-test. For the qualitative data, a semi-structured interview form developed by the researcher was used and students in the experimental group were asked about the flipped course. During 16-week experiment, the experimental group was asked to watch videos about the lecture content and complete the short quizzes at home before the course in the classroom. The classroom activities were learner-centered to facilitate classroom interaction. On the other hand, the students in the control group were instructed following classic blended learning method. The collected quantitative data was analyzed using t-test, one way Analysis of Covariance (ANCOVA), and Multivariate Analysis of Variance (MANOVA) while descriptive analysis was used to analyze the qualitative data based on pre-determined themes. As a result, it was explored that the flipped model did not exert a significant impact on two groups of students' weekly performances. However, supporting its role in the retention skills, students taking the flipped course were found to be significantly more successful than the others in the control group based on the post-test scores. In terms of self-directed learning readiness total scores, the students did not differ. Yet, the flipped model was found to be effective on students' motivation scores. In addition to its positive impact on students' learning motivation, the qualitative revealed the model also fostered positive opinions about the course. Being mostly compatible to the results covered in the literature, Sırakaya (2015) reflected the model as a tool to help the learners develop a prejudice-free attitude toward the course.

Another recent study was utilized by Yavuz (2016). This study was reviewed in this part since it, on the contrary to other cited thesis studies, was conducted with the participation of 27 female students at a Vocational and Technical High School in Mardin. This study followed a mixed method design to find the impact of the model on students' academic performance and their perspectives of the model. In this respect, ten graders followed their Information Technology course following a flipped model for four weeks in 2014-2015 academic year. The data was acquired through the administration of a multiple-choice achievement test developed by the researcher. Students' opinions about the course were elicited from focus group

interviews following each flipped course during four weeks. The quantitative data collected in the quasi experimental design was analyzed using *Independent Sample t-test*. On the other hand, the qualitative data was analyzed following content analysis approach. In the first fold, the study yielded results contradicting with most of the literature. That is to say no significant difference was found between two groups in terms of their scores of the achievement test. On the other hand, the focus group interview produced comprehensive data about the negative and positive aspects of the model from students' perspectives. In this vein, students reported that the flipped classroom model could be broadened to other disciplines since it increased their motivation to learn. However, the participant students also provided the instructors and designers of the flipped classroom model by warning them about necessity of setting up technological infrastructure and training program about the model both for teachers and students.

In addition to seven thesis studies (Akgün, 2015; Aydın, 2016; Balıkçı, 2015; Gençer, 2015; Sırakaya, 2015; Turan, 2015; Yavuz, 2016) conducted by researchers from the computer related departments, there are two more thesis (Kara, 2016; Yiğit, 2014) focusing on flipped classroom in non-EFL fields in Turkey. The reason behind reviewing these two studies in a different vein is that these studies were performed by researchers majoring in fields that are neither computer related nor non-EFL. In order to expand the scope of the review of the present study, these two studies were also addressed briefly below.

As the first of these aforementioned studies, Yiğit (2014) contributed to the diversity of the studies focusing on flipped classroom model. Following a phenomenological approach, Yiğit (2014) designed a case study to evaluate the model based on the five decision stages of *Rogers' Diffusion of Innovation Theory*. In this vein, the case of the study was adopted a private college in Istanbul using the flipped classroom model in some courses. In the fall and spring terms of 2013–2014 academic year, Yiğit (2014) aimed to collect data from all the stakeholders of the model including 2 administrators, 17 teachers from different courses ranging from Social Sciences and English to Science and Maths, 17 students and 4 parents. In this respect, all the factors effective in the diffusion of the flipped classroom model were aimed to be covered through semi-structured interviews, observation and document analysis. Different questions were asked to different stakeholders in the interview protocols.

One of two social sciences classroom where the model was implemented was observed to see it in practice. As documents, forms developed by the instructors in the flipped classrooms were collected to support the results acquired from the data. Descriptive and content analysis were performed to analyze the data. As a result of the analysis, for all the stakeholders, the model was found to be highly related to video lectures in online environments and tablet based learning. Within *Diffusion of Innovation Theory,* the necessity was effective on the decision to implement the model while students' age groups, grades and their obligation to study for a central exam were found to be reasons behind not using the model in the classrooms. In this doctorate study where the flipped classroom model was adopted as not only a phenomenon but also an innovation, the model yielded results mostly compatible to the related literature. However, the positive and modificative impact of the model on students' and their parents' learning culture was observed to be limited to the students for whom and the classroom where the courses were flipped. In this vein, through the engagement of a wide range of stakeholders to the research process, Yigit (2014) reflected the flipped classroom as beneficial for the students. However, it was found to be complex for the teachers requiring a great deal of preparation before the class and experience to apply it and a flexible learning model requiring the support of the school management. With its different research design, by providing a mirror at the dynamics hindering the model in Turkey, the study promises valuable information to the concerned bodies who aim to spread the model to a number of disciplines.

The latest and second of the aforementioned two thesis studies, centered on the fields of study related to neither computer education and instructional technology nor EFL, was an attempt to reflect flipped classroom model into medical education. In a similar way to the studies reviewed above, Kara (2016) designed a mixed method master thesis study to evaluate flipped clinical trainings at *Otorhinolaryngology Department,* in Pamukkale University Medical Faculty, in 2014-2015 academic year. This evaluation included the impact of the model on undergraduate medical students' (n=127) academic achievement, their satisfaction and opinions of their flipped classroom experience. However, it is important to note that the study did not assign the participants to control or experimental groups. The participant students were asked to take an achievement test developed by the researcher as a pre and post-test.

Additionally, these students were asked to complete a survey developed by the researcher to measure their opinions about and satisfaction of flipped experiences. The survey was reported to be a 5 point Likert scale including some open-ended items. As the last source of the data, students were asked to write feedbacks at the end of clinical training. The results of the achievement test indicated the significant difference between pre-test and post-test scores. This supported the view in the literature that the model has a positive impact on academic performance of the learners. The results of analysis of the data acquired from the surveys revealed that over 80 % of the students were found to be satisfied with the flipped classroom model and 96 % of them learned the content deeper. Moreover, over 90 % of the students were reported to meet the cognitive objectives of the course. On the other hand, content analysis of the qualitative written feedbacks explored that students defined the flipped classroom experience with a focus on the word *efficient.* Providing them better options to practice the theoretical content by meeting with more patients in the flipped clinical trainings, the flipped classroom model was recommended by the participant students to be applied in other clinical trainings. In brief, in this study, Kara (2016) produced results that prove a wide range of advantages of the flipped classroom model in medical education.

Similar to the limited number of articles taking stock of flipped classroom model in Turkish EFL setting, the search for "flipped classroom" in the database of *Council of Higher Education Thesis Centre* found out only 2 (Boyraz, 2014; Ekmekçi, 2014) of 11 studies focused on flipping EFL courses in Turkey. One of them was a doctorate thesis (Ekmekçi, 2014) seeking to evaluate the model in EFL writing. The other study, on the other hand, was a master thesis (Boyraz, 2014) setting out to assess the role of the model in EFL learners' academic performance in the acquisition of two grammar structures. To provide the basis for a direct comparison of the results produced by present study with those reached in similar settings, these two studies were briefly reviewed.

In the first one of the aforementioned studies, Ekmekçi (2014) aimed to propose flipped classroom model with a focus on blended learning for EFL writing classes. Hence, the study included 43 university students enrolled ELT prep classes in *School of Foreign Languages* at *Ondokuz Mayıs University*. In this study, Ekmekçi (2014) followed a mixed method design employing both qualitative and quantitative data

collection instruments. Accordingly, pre- and post-test true experimental design constituted the quantitative part while the semi controlled interviews stood for the qualitative part. In the fall term of 2013-2014 academic year, the treatment in the experimental group lasted 15 weeks. This group of learners was instructed following Flipped Writing Class Model. In this respect, 23 students in the experimental group were provided with video lectures created by the researcher or videos from other sources. Through a course management system called *Edmodo,* students reached and watched these videos at home as a result of which they were engaged with more exercises and more time was shared for paragraph writing in the classroom. On the other hand, 20 students in the control group were instructed following a traditional lecture based approach. These two groups of students were instructed by the researcher based on the same curriculum and book. To collect quantitative data, the pre- and post-tests, *Argumentative Paragraph Rubric*, *ICT literacy survey*, and *Flipped Writing Class Attitude Questionnaire* were administered. *Argumentative Paragraph Rubric* developed by the researcher was used to measure students' performance in writing an argumentative paragraph on one of three topics given by the researcher. This procedure was realized as both pre and posttest. ICT literacy survey was employed to the students in the experimental group after the pre-tests to evaluate their tendencies, including their equipments such as internet or computer at home, to use the model. The 5-point Likert *Flipped Writing Class Attitude Questionnaire* consisting 25 items were developed to measure the students' attitudes toward flipped writing courses (α= 0,92) to support the data produced by the questionnaire, a semi-controlled interviews including 7 questions were developed and applied to the students in the experimental group. As a result of the analysis of the collected data, it was understood that the students in the flipped classroom were significantly more successful than those in the traditional classes. The results of ICT survey indicated that students had the necessary technological infrastructure at home and sufficient knowledge to use it. In terms of students' attitudes toward flipped writing class, the study yielded positive results. Supported by the qualitative data, the students were found to think that flipped writing class was more enjoyable and effective than traditional classrooms. In brief, Ekmekçi (2014), by focusing on only one skill area, explored positive results compatible to those covered in the related literature and ascertained the flipped classroom model as an efficient way of improving EFL undergraduate Turkish learners' writing skills.

The second and the last thesis study on flipped classroom model in EFL in Turkey was utilized by Boyraz (2014). With some limitations, the study was performed following the similar research design seen in the related literature. Accordingly, the study was designed based on a mixed method. In Pre-test post-test control group experimental design, Boyraz (2014) aimed to explore the impact of the flipped model on the undergraduate EFL learners' academic performance on the acquisition of two grammatical structures and its retention. On the other hand, standing for the qualitative part of the study, it was attempted to reveal students' opinions about the model through focus group interviews. 42 students in two different classes from English preparation class at Aksaray University were participated in the study in the spring term of 2013-2014 academic year. Each group participated both as an experiment and control group in the study. In the control group, the researcher employed traditional approach based on lecture presentation, question and answer in the classroom. The rest of the time was allocated for the exercises in the coursebook. The participants in the experimental group, on the other hand, watched the video sent them through *Edmodo* before the course at home and came to classroom having finished short exams to ensure that they watched the videos. In this way, this group of EFL learners was able to spend more time in the classroom on the acquisition of higher order skills related to the grammatical structures within the scope of the study. The researcher developed two achievement tests covering two grammatical structures, indirect speech and passive voice, and used them both as pre- and post-test. For students' opinions about the model, the researcher developed an interview form including 8 questions. 34 students participated in the focus group interviews and based on 8 themes standing for each of the research questions, the qualitative data was analyzed in two groups as negative and positive following content analysis method. The achievement tests were analyzed through *independent* and *paired sample t tests*, *Mann-Whitney U* and *Wilcoxon test* depending on the distribution of the data. The results of the data revealed the significant difference in the scores of two groups of learners in the favour of those taking the flipped courses. The results of the interviews, on the other hand, explored that 73,77 % of the participant students held positive opinions about different dimensions of the flipped model. As a significant contribution of the study to the literature, Boyraz (2014) indicated the long-term positive impact of the flipped classroom model on EFL learners' academic performance.

To sum up, the review of literature on flipped classroom in both foreign and Turkish setting does not reflect the attention it has recently attracted. This was proved with 82 studies found in the database *web of science*, 4 articles seen in the database of *Turkish Journal Park Academic* and 11 thesis studies explored in the database of *Council of Higher Education Thesis Centre.* These 97 studies were demonstrated to be published from 2012 to 2016 with a focus on flipped classroom in different educational fields at different educational levels through the world. Furthermore, this detailed evaluation of the related literature ascertained that, on a large scale, research in STEM fields gave birth to the flipped classroom model while its application in flipped EFL or ESL setting is rarely seen (Engin, 2014). Based on the number of the flipped studies in EFL or ESL settings, the picture is better and easily viewed. Accordingly, the review above showed 7 of 82 studies from *web of science*, 2 of 4 articles from *Turkish Journal Park Academic* and 2 of 11 theses from *Council of Higher Education Thesis Centre* were carried out to locate the flipped model in the different dimension of EFL. Therefore, to the interest of the present study, this indicates totally 11 flipped studies, 4 of which were done in Turkey. Hung (2015) elaborated on this by stating "in the field of language education, little or no research to date has rigorously studied whether and how flipping the language classroom can enhance student learning" (83). Considering the advantages of the model cited in the related literature, it seems surprising to come across with a few studies focusing on flipped classroom model in EFL or ESL (Leis, Cookie and Tohei, 2015).

Based on the review of the studies in some databases, there appeared rare attempts to reflect the flipped classroom into different dimension of EFL at higher education in Turkey. Through either qualitative or mixed method design, it was aimed to shed lights on EFL undergraduate learners' opinions about, attitudes toward the flipped classroom model and its impact on their academic performance. Their performances in flipped grammar, reading or writing courses were separately measured in these studies. Sharing lots of common with the research results in the international scope, it was explored that flipped classroom served as facilitator of both EFL performance and attitudes toward EFL in Turkey by providing "action based, authentic, connected and collaborative, innovative, high level, engaging, experience based, project based, inquiry based, and self-actualizing" activities in the classroom (Hamdan et al., 2013, 17). On the other hand, a bulk of studies in the literature has also consistently

revealed self-efficacy and attitude as the predictor of achievement in EFL in Turkish setting. (Başaran, Cabaroğlu, 2015; İlhan, Karatas, 2015). Despite the cited significant roles of self-efficacy belief, attitude and flipped classroom model in EFL, more research is suggested to form a uniformed basis on their potentials in EFL (Bishop, Verleger, 2013; Yurtseven, Altun, Aydin, 2015).

2.2. Attitudes of EFL Learners toward English

Language is a significant means of transferring one's own thoughts, feelings, joys, worries and simply oneself (Tavil, 2009). The complex nature of rapidly globalizing world initiates the arguments regarding the use of English as an international language (İnceçay, Aysel, 2014). This has recently brought the assertion that English 'belongs to everyone who speaks it, but it is nobody's mother tongue' has been fore fronted (Rajagopalan, 2004). As a result of this, in order to facilitate the communication among the non-native members of the big society, the global world, another argument is that English should be taught not as a native language within the boundaries of Standard English, but as a language of the all with a new approach tolerating accent variations and some structural mistakes while diminishing the native authority on the language (Jenkins, 2008).

Considering the significant functions, it realizes in today's world, EFL has been a research area of growing interest in the literature. Research into the use of English in foreign language context has shed lights on the impact of some personal factors such as motivation, attitudes, anxiety, aptitudes, intelligence, age, gender, personalities etc on EFL learners' performance during the learning process (Gardner, 1985). Among these cited factors, EFL learners' attitudes toward the language they are to learn is discussed as one of the factors that exerts most impact on the achievement in language learning process (Fakeye, 2010; Obeidat, 2005).

Within this in mind, the current study delves into the impact of flipped classroom model application, the intervention in the experiment model, on EFL students' attitudes toward EFL. Therefore, the following part will be devoted to explore the perception of English as a second language in a foreign context covering the definition of attitude including its behavioral, emotional and cognitive aspects, its role in achievement in EFL revealed by related studies.

2.2.1. Definition of Attitude

It is widely accepted in the related literature that students' attitudes toward English are playing a significant role in their success in the process of learning it as a foreign language (Tsiplakides, Keramida, 2010). Owing to its function as the predictor of success in varying fields in education in addition to EFL, a number of researchers in the fields of psychology and education have touched on the topic. As a result, several definitions of attitude including different meanings from different contexts and perspectives were formed in the literature (Abidin, Pour-Mohammadi, Alzwari, 2012).

As a construct that cannot be observed directly but can be inferred from individuals' behaviors' or their self- reports, attitude is simply defined by Montano and Kasprzyk (2008, 71) to include "individual's beliefs about outcomes or attributes of performing the behavior". Similarly, Gardner (1985) asserts that attitude is an individual's evaluative reaction to an object or a referent, which can be inferred from that individual's beliefs or ideas about the referent. These definitions emphasize the close relationship between attitude and the individuals' behavior. As a result, it is possible to claim that attitude is not simply related to the outputs of one's behaviors but also one's values, beliefs and opinions that underpin and shape his or her behaviors. Therefore, a person who has strong and positive values and opinions about the object of the behavior will form positively valued outcomes while a person who holds strong beliefs that negative values and opinions about the referent of the behavior will tend to form negatively valued outcomes.

Close relation between the behavior and the attitude made it necessary to make a broader definition of the term in a way to cover all dimensions of a behavior. Consequently, on the contrary to the aforementioned simple definitions of the term, attitude, a pioneering concept in social psychology for a long time, the definitions of attitude reviewed in the related literature were explored to encompass cognitive, affective and behavioral dimensions. Covering these three components, Bohner and Schwarz (2001) reflect the composition of those definitions in the literature as "an enduring organization of motivational, emotional, perceptual, and cognitive processes with respect to some aspect of the individual's world" (437).

In the current study, the focus of the study is on the impact of flipped classroom model on EFL students' attitudes toward English. Within this respect, the referent of the attitude, in other words, the behavior is the language itself. As a result of the nature of the behavior researched in the current study, the definition of the attitude was enlarged to include attitude toward a language and learning that language. Thus, definition of 'language attitudes' in the Longman Dictionary of Applied Linguistics (1992) was adapted in this study. It was, consequently, defined, with an emphasis on its function as the predictor of language performance as follow:

> "The attitude which speakers of different languages or language varieties have toward each other's' languages or their own language. Expressions of positive or negative feelings toward a language may reflect impressions of linguistic difficulty or simplicity, ease or difficulty of learning, degree of importance, elegance, social status, etc. Attitudes toward a language may also show what people feel about the speakers of that language" (191).

Compatible with the broad nature of attitudes, language learning process has intellectual psychological and social loads (Abidin, Pour-Mohammadi, Alzwari, 2012). Correspondingly, Gardner and Lambert (1972) state that students' ability to learn a second or a foreign language is influenced by not only their intellectual capacity or linguistic competence but also their attitudes toward that language. As a result, three-dimensional nature of attitudes, a significant factor in language learning process including the changes in emotional, psychomotor and cognitive domains, has the potential to facilitate or hinder the language learning process (Clément, Dornyei, Noels, 1994).

Within this respect, the concept of attitude in the current study was viewed from Wenden's (1991) comprehensive presentation of attitudes. Accordingly, as seen in the following figure below, Wenden (1991) regarded the construction of attitudes as a composite of three interrelated domains: behavioral, cognitive and affective. These three aspects of the concept of attitude, theoretical frameworks of whom were shaped by approaches of behaviorism, cognitivism and humanism respectively, were briefly described in the following part.
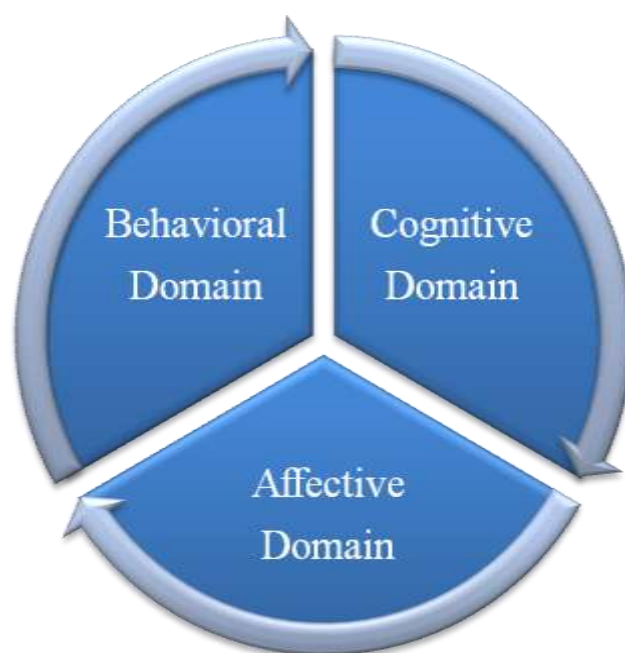
Figure 2.1: Construction of Attitudes

2.2.1.1. Behavioral Domain of Attitude: In terms of learning a foreign language, behavioral aspect of attitude stands for students' tendency to behave and react in particular learning contexts. This domain reflects that it is possible to have an idea about one's attitude toward an object from his or her behaviors. In EFL context, this leads to form a basis on which judgments about language learners' attitude toward the language they study can be easily made. Drawing on this basis, researchers may infer that students enjoy learning a language when they are deeply engaged in language learning activities without the drive of any kind of pressure or reward. On the other hand, it may be concluded that they do not like spending time on learning a language if the learners are made to engage in the activities just for the sake of an external reward or avoiding a punishment.

Within this respect, as asserted by Jones (1979), by following the same rules, an observer can make inferences about students' attitudes toward the language they study from their related behaviors.

To conclude, when the cited relation between successful language learning and attitudes toward the target language is taken into account (Gardner and Lambert, 1972), it becomes significant to be familiar with the language behaviors as the predictor of positive attitudes. Kara (2009) elaborates on this with the examples of

behaviors that can be performed by successful language learners holding positive attitudes toward their language of study as follow:

> "Positive attitudes led to the exhibition of positive behaviors toward courses of study, with participants absorbing themselves in courses and striving to learn more. Such students were also observed to be more eager to solve problems, to acquire the information and skills useful for daily life and to engage themselves emotionally, thereby meeting the requirements of the courses in terms of behaviors, emotions and psychomotor skills" (102).

2.2.1.2. Cognitive Domain of Attitude: Cognitive dimension of attitude involves the beliefs of the language learners about information processing dominant in foreign language learning context. In other words, not only learners' actual behavior but also their perceptions of and beliefs about those behaviors shape judgments of attitudes. Therefore, language learners' beliefs and perceptions with regard to the object of study that construct the basis of behaviors are emphasized. Within this respect, this component reflects the fact that in EFL contexts, learners have little contact with native speakers while they have a great deal of indirect contact with target culture thanks to the versatile means of media (Dörnyei, 1998).

In brief, the cognitive attitude comprises following four steps of information processing in foreign language learning: "Connecting the previous knowledge and the new one, creating new knowledge, checking new knowledge, and applying the new knowledge in many situations" Abidin, Pour-Mohammadi, Alzwari, 2012, 122). In addition to its role to underlie the implicit components of foreign language learning process, the aforementioned part on cognitive aspect of the attitudes also illuminates cognitive load of attitudes and significant relationship with the behaviors. This nature of the attitudes, accordingly, urges the researchers to delve into the concept of cognition in foreign language learning process.

2.2.1.3. Affective Domain of Attitude: Affective aspect of language attitudes encompasses emotional loadings in foreign language learning process. In other words, this component refers to learners' feelings about objects, people or situations. Within the current study, it stands for EFL learners' likes and dislikes about the target language they are supposed to learn. Therefore, this domain of attitudes is also labelled as evaluative component in the related literature defining attitude with an emphasis on emotions as "a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavour (Bohner, Schwarz, 2001, 437).

Evaluative nature of the concept of attitude implies that language learning is an affective process that includes not only emotional traits peculiar to learners but also shaped by them. Besides simply indicated as likes and dislikes, some of these traits are also exemplified by Dörnyei (1998) as confidence, fear, anxiety and self-efficacy. In this respect, it is simply agreed that due to the influence of such factors, an individual learner can like or dislike to study and learn a foreign language. However, this affective reaction to that language exerts a significant impact on their language performance (İlhan, Karataş, 2015).

Within the perspective of affective boundaries of attitudes, it is important to note the assertion made by Dörnyei (1998) that an EFL learner's ability to succeed in language learning process is significantly related to his or her emotional readiness to learn that language. As a result of this, it is important to take the psychological and emotional aspects of foreign language learning process into consideration.

To conclude the whole concept of attitude, it can be stated that these aforementioned aspects of attitude imply EFL learners' language attitude is multi-dimensional. That is to say, when taking a step in the process of language learning, students can draw on different sources of information to make an evaluation and perform a behavior accordingly. While this process of making an evaluative judgment and performing an output can depend on learners' own feelings or their own perceptions of the target language, it can also be fed from their own learning experiences.

Compatible with the framework drawn by three folded nature of attitudes, Bloom (1995) suggests that learners need to be at specific required cognitive and affective levels to perform successful behaviors in the process of learning. Thus, without heavily relying on only one limited source of positive attitude restricted to only observed behaviors of the learners, it is highly important for the researchers to internalize the understanding that language attitudes embrace behavioral, cognitive and affective aspects, which has a close relationship with successful language learning. Therefore, in the current study, the researcher delved into the concept of attitude toward learning EFL as a multi-dimensional construct relying on varying sources of information.

2.2.2. Significance of Attitudes in EFL

Significance of attitudes in foreign language learning lies in its cited impact on the level of success attained by the learners. This fact is not only supported by a number of researches carried out in the related literature but also the definition of learning itself. Within this respect, as a composition of different theoretical views of learning, Kara (2009) defines learning as the process of acquiring the skill to adapt oneself to the constantly changing conditions in the environment relying on a positive change in the cognitive, emotional and behavioral domains. In the concept of EFL, this composite view of learning implies that learning a language should not be taken as an academic entity deeply based on such steps of cognitive processes as relating the previous knowledge to the new knowledge, applying new knowledge in different real life situations, analyzing and accordingly synthesizing new knowledge. On the contrary, due to cultural, political and affective loads of a foreign language, EFL needs to be "approached primarily as a social and psychological phenomenon" (Abidin, Pour-Mohammadi, Alzwari, 2012, 121). This strongly explored reality in the related literature motivates the researchers to accept that level of success in EFL is not only related to learners' intellectual or cognitive capacity but also to their self-reported or observed attitudes toward the target language and learning that language itself (Gardner, Lambert, 1972). On this point, attitudes are said to be closely related to cognitive, social and psychological with a multifaceted nature and explored by a number of studies in national and international scope. These studies together set attitudes as one of significant key factors predicting EFL learners' language performance (Abidin, Pour-Mohammadi, Alzwari, 2012; Dörnyei, 1998; İlhan, Karataş, 2015; Gardner, Lambert, 1972; Kara, 2009).

In addition to its role as the predictor of success in EFL, there is also a view that sees attitudes as the component of motivation (Gardner, Lambert, 1972). Its predictive power of motivation adds to its precious place in EFL. This is because of the fact that motivation is regarded by both teachers and researchers as a phenomenon determining most of the rate of success or failure in both second and foreign language learning process. As a researcher spending remarkable amount of time on it, Dörnyei (1998) clearly enlightens the role of motivation in second/foreign language learning as follows:

> "Motivation provides the primary impetus to initiate learning the L2 and later the driving force to sustain the long and often tedious learning process; indeed, all the other factors involved in L2 acquisition presuppose motivation to some extent. Without sufficient motivation, even individuals with the most remarkable abilities cannot accomplish long-term goals, and neither are appropriate curricula and good teaching enough on their own to ensure student achievement. On the other hand, high motivation can make up for considerable deficiencies both in one's language aptitude and learning conditions" (117).

As a result of this explanation, it becomes clear that high level of positive motivation leads to higher academic performance in EFL. Attitude, in addition to its direct relation to language achievement, is an important affective factor exerting an impact on motivation.

To conclude, attitude is accepted as a significant factor in foreign language learning process. As set by Csizér and Dörnyei (2005), attitude influences learners' choice of the language they will study and the amount of time they will devote to learning that language. Therefore, an individual holding positive attitudes toward a subject area gets interested in it and, accordingly, bears all its problems enthusiastically during the process. On the other hand, as suggested by Gardner (2006), in the form of chain, positive attitude toward language learning leads to high level of motivation and inevitably better performance in EFL.

2.2.3. Research on Attitudes toward EFL

The related literature provides a considerable amount of studies focusing on different aspects of attitudes toward learning a foreign language in general and English as a second or foreign language in particular and its impact on language learners' learning behaviors and performance (Abidin, Pour-Mohammadi, Alzwari, 2012; Kara, 2009; Obeidat, 2005). In specific terms, in addition to academic achievement, these studies attempted to explore the impact of attitudes toward the language and its aspects including ideology, people and culture on a number of variables such as motivation, learning strategies, language learning at different academic levels, non-native speakers, language teachers (Karahan, 2007). However, as declared by Obeidat (2005), an outstanding amount of these studies were conducted to shed lights on the EFL learners' attitudes toward English as a foreign or second language and its effect on language learning process. Within this respect, in order to form a basis for discussing the results generated by the current study, some recent studies together with some fundamental ones focusing on students' attitudes toward EFL were covered below.

Pioneering studies on attitudes toward learning a foreign language in general are Csizér and Dörnyei (2005), Dörnyei and Ottó (1998) and Gardner (2006). While Csizér and Dörnyei (2005) set, as a result of their study, attitude as a predictor of success in learning a foreign language, Gardner (1985) indicated the indirect impact of positive attitudes toward learning a language on the motivation to learn that language. On the other hand, Dörnyei and Ottó (1998) in their Process Model of L2 Motivation detailed this impact. They (1998), in this regard, explored its effect on the choice of the language to be studied and the amount of effort devoted to learning it.

In the wide scope of international studies, the significant relation between attitudes and language learning performance is frequently highlighted (Dörnyei, 1998; Karahan, 2007). Sharing the same basis as Visser (2008) who explored, in a study focusing on the case of German Army, that students' positive cognitive attitudes toward knowledge acquisition enhanced learning process, Fakeye (2010) followed the similar trace in EFL in African context. In a study conducted with 400 Nigerian secondary school students studying EFL, Fakeye (2010) revealed the significant relation between attitude and language learning performance. Exploring no gender impact on attitude, Fakeye (2010) indicated those holding positive attitudes toward EFL as better EFL learners.

As an example of studies on attitudes toward EFL in Arabic context, Momani (2009) led a research on Jordanian High School students' attitudes toward EFL and their performance in EFL reading. The results supported the highly cited relation between positive attitudes toward EFL and better academic performance.

As a representative of a bulk of studies on attitudes toward EFL in Asian context, Rukh's (2014) study on business students in Punjab was adopted. With the participation of 200 business students, Rukh (2014) attempted to shed lights on the relation between the business students' attitudes toward EFL and their academic performance in it. The results of the study obtained from the administration of a 3-point Likert scale concluded the business students' positive attitudes toward EFL and its relation to their academic success.

As in the rest of the world, the predictors of success in EFL, which is becoming more related to globalization day by day, are questioned in the Far East context. Within this perspective, the study carried out by Liu (2007) represents the interest in the

impact of affective factors on EFL performance. In her study, Liu (2007) administered a 44-item survey adapted from Gardner's (1985) to 202 third-year EFL students in a southern university in China. The results of this study supported Gardner's (1985) assertion that attitudes toward a language exert impact on motivation to learn that language and indirectly language performance. In other words, Liu (2007) explored that Chinese EFL students held positive attitudes toward learning English and motivated to learn it, which was significantly associated with their English proficiency.

In the Latin American context, there have been a number of studies focusing on attitudes of students, parents or teachers toward learning English. Representing those efforts, Martínez, Pérez and Madrid (2013) studied on 110 Mexican American students' attitudes toward learning English as a second language in a structured immersion program. They (2013), as a result of the study, found attitudes toward learning English including many aspects such as attitudes toward teacher, teaching strategies employed, materials and tasks used in the classrooms classes as the main predictor of motivation and success.

In expanding circle countries, English is taught as a foreign language with the aim of more qualified international communication (Çetinkaya, 2009). European countries are clear examples of this circle. In these countries similar to those in the outer circle where English is a compulsory second language, there is a great deal of effort to understand the role of attitudes toward EFL in language performance. Within this respect, a study conducted by Liliequist (2013) embodies these efforts. In the study, it was aimed to find the difference between secondary school high and low achievers with respect to their motivation and attitudes toward EFL. 86 students were included in the study. They were identified as low and high achievers on their grades at school. As a result of the study, Liliequist (2013) revealed that high achievers held significantly more positive attitudes toward EFL with a better interest in it than low achievers.

In addition to those studies where English is adopted as a foreign language, there are also studies conducted on the attitudes toward learning English in outer circle countries such as India, Kenya, Malaysia, Singapore, and Zambia. Outer circle, by the way, "refers to the regions where English is used as an institutionalized second language as a result of colonialism" (Çetinkaya, 2009, 110). As one of those studies

in this circle, Latifah et al. (2011) conducted a nationwide research with directed 757 university learners. In the study, the researchers aimed to explore the impact of a number of affective factors such as motivation, attitude, and anxiety on their performance in English language course in Open University Malaysia. The results obtained from the study indicated the positive correlation between aforementioned affective factors with language learners' performance. In this respect, the wide-scope study explored that together with motivation attitudes led to better performance in EFL.

In contrast to a bulk of studies in international scope, Çetinkaya (2009) complains about the limited number of studies on EFL Turkish students' attitudes toward learning English in the national context. Furthermore, she explains that most of those studies followed a quantitative design employing structured questionnaires to explore language learners' attitude. In this part of the study, to provide the melting pot of the results obtained from international and national research, some studies were presented as the representatives of those entire carried out in Turkish context.

In an older study, Gömleksiz (2010) attempted to take a full photo of the concept of attitudes toward EFL in Turkey embracing the probable impact of such variables as gender, grade level and department. With 1275 students at Fırat University in Elazığ in Turkey, Gömleksiz (2010) revealed that students' attitudes toward EFL differed significantly in terms of their gender, grade level and department. In contrast to other studies mentioned above, this study indicated that female students had more positive attitudes toward EFL than males. In addition to the departments they study, students' attitudes also differed in terms of their grade levels. In other words, sophomores' attitudes were more positive than a freshman, which indicates that spending more time on the language and being academically more successful are related to positive attitudes toward EFL.

One of the most comprehensive studies in Turkish context was carried out by İlhan and Karataş (2015). With 447 university students learning EFL in Istanbul were included in the study. The researchers aimed to gain insights about the EFL students motivational beliefs, attitudes regarding learning English in Turkey. As a result of this study, no gender impact on neither motivation nor attitudes was found. Moreover, another fold of the result supported Gardner's (1985) view on attitudes

that links it to motivation. In other words, EFL students' motivational strategies were explored to be significantly correlated to their attitudes toward learning English.

Another recent study conducted by Şentürk (2015) with 61 university students learning EFL at Zonguldak Bülent Ecevit University enlightened the role of attitudes in a different aspect of EFL. By employing Foreign Language Reading Attitudes and Motivation Questionnaire to 61 intermediate, pre-intermediate and beginner level EFL learners, set by the results of placement test they took at the beginning of the term, this study examined the relation between EFL learners' attitudes toward EFL reading and their performance in it. While the results indicated no gender impact on their attitudes toward EFL, significant differences were found among learners in terms of their proficiency levels. That is to say, intermediate level EFL readers had more positive attitudes toward EFL reading than pre-intermediate and beginner level EFL learners.

As a result, this part covered two folds of studies on attitude toward EFL. In the first fold, studies conducted in the countries representing nearly all parts of the world where language is taught as a compulsory second language and as a foreign language in the international scope were covered. In the second fold, three recent studies conducted with university students in different regions of Turkey were reviewed. As a composition of these studies in the national and international scope, the literature indicates attitude as a significant predictor of language achievement. This body of research also suggests EFL students' positive attitudes toward the language they will study are significantly related to motivation to learn the language leading to the improvement in the quality of their learning. More specifically and significantly, Gardner and Lambert (1972) suggested that an overall understanding of students' attitudes will contribute to curriculum developers and designers to tailor language teaching programs in a way to tap the students' language learning attitudes and motivation as the key of having more successful EFL learners.

In contrast to frequently cited roles of attitudes in EFL in international scope, the current research sets to fill in the gap in the national scope. As suggested by Yurtseven, Altun, Aydin (2015), on the contrary to substantial body of research producing similar results in the international scope, studies carried out with university students on language motivation and attitudes in Turkey seem not satisfactory to form such a uniformed basis. Furthermore, touching on the role of

English in EFL Turkish students' personal and professional lives, as suggested by Çetinkaya (2009), quantitative nature of studies on attitudes toward EFL should be abandoned to better understand the concept of attitude in EFL.

Taking these complaints about the gap in the related Turkish literature into consideration, the current study is based on the probability put forward by Abeysekera and Dawson (2015) that flipped classroom model may help students get more motivated to spend a considerable amount of time on learning. This will, in turn, "provide special opportunities to make this work more manageable, achievable and tailored to each student through the management of cognitive load" (8). Within this respect, adopting a mixed method, the current study aims to delve into the impact of flipped classroom model on EFL students' attitudes toward EFL, as a well accepted significant predictor of EFL performance.

2.3. Self-Efficacy Beliefs of EFL Learners in English

With the rise of humanistic approach in 1960s, the tendency to explain everything in educational contexts referring to cognition was abandoned (Gardner, 1985). Since then, affect in addition to cognition was considered as an important determinant of academic achievement in different fields. As a result of this new tendency, it has been emphasized by many scholars in EFL that studies focusing on explaining the learners' different language learning performances with affective factors must be carried out. Within this respect, affect has been regarded as a significant factor exerting its impact on learners' academic achievement in foreign or second language learning process (Gardner, Lambert, 1972).

Affect is the composite of a set of variables predictive of academic performance in different fields (Dörnyei, 1998). These variables are many in number ranging from learners' needs, expectations to their attitudes toward, interest in the subject matter and their beliefs in their ability to be successful. These refer to different components of motivational orientations. Among these variables, learners' perception toward learning English as a foreign language and their beliefs in their abilities to be successful language learners are the focus of the current study. Within this perspective, in addition to the previous part on attitudes, this part will be devoted to shed lights on *self-efficacy* as a significant component of motivational orientations

predicting language performance covering its definition and review of the related studies.

2.3.1. Definition of Self-Efficacy

Recently, there has been an influx of interest in self-efficacy as a significant psychological construct having a predictive power of success in varying cognitive processes including foreign language learning (Başaran, Cabaroğlu, 2015). The theoretical foundation of this construct is based on Bandura's social cognitive theory. In this theory, human behavior is simply explained to be determined as the result of reciprocal relation among the individual, the environment surrounding the individual and the behavior itself (Bandura, 1978). Having an important part in this reciprocal relationship, self-efficacy is defined by Bandura (1997) as "beliefs in one's capability to organize and execute the courses of action required to manage prospective situations" (2). In this respect, self-efficacy can be regarded as one of the arteries of Bandura's social cognitive theory since people's beliefs about their ability to accomplish a given task are decisive in not only their social bounds but also cognitive outputs including their devotion to the goal attainment, fight with anxiety, stress or depress.

In the related literature, self- efficacy is frequently referred as one of the significant chains that takes the learner to positive motivation and accordingly to the ultimate achievement. Dörnyei and Otto (1998), stressing its role in the goal formation and attainment process, set self-efficacy as the sum of interactive factors determinant in self-evaluation process. They elaborate on it as follow:

> "A selected goal, by definition, must have already passed the test of potency, that is, it must have been regarded as broadly attainable. In the intention formation phase the expectancy of success is more specifically assessed. Based on a number of interacting factors such as self-efficacy/self-confidence, perceived goal difficulty, the amount of expected support, L2 anxiety, perceived competence, the quality and quantity of previous L2 contact, and causal attributions about past experiences (successes and failures), the individual makes an evaluation of his/her *coping potential* in the planned action. The greater the perceived likelihood of goal attainment, the higher the degree of the individual's positive motivation. Conversely, it is unlikely that effort will be invested in a task if the individual is convinced that he/she cannot succeed no matter how hard he/she tries" (54).

In addition to its widely accepted definition as one's perceived ability to succeed in a given subject matter or a task, it is also necessary to touch on the matter of incomprehensibility (Başaran, Cabaroğlu, 2015). Accordingly, in the related literature some researchers frequently stress the word "belief" when they refer to the

concept while some others use the word "perception". As a result of this interchangeable use of the term, some phrases such as "self-efficacy beliefs" or "self-efficacy perceptions" or "perceived self-efficacy" appear in the literature. Sharing the same basis as Başaran and Cabaroğlu (2015), the term "self-efficacy beliefs" was adopted in the current study. The reason behind this choice is expressed as follows:

> "The term "belief" seems to better reflect the complex nature of the term, for, unlike perception, it connotes that self-efficacy has deep roots in past experiences and close mutual ties with not only context but also personal traits and psychological constructs such as passion, tenacity, motivation and anxiety" (50).

When the posited relation between self-efficacy beliefs and goal attainment is taken into consideration, it seems as a must to better conceptualize the sources of those beliefs. Regarding this, Bandura (1997) asserted that four different sources drove people's beliefs about their capability to accomplish a task. These are mastery experiences, vicarious experiences, social persuasion and psychological arousal. These basics of the self-efficacy beliefs were briefly provided below.

2.3.1.1. Mastery Experiences: The most influential source of creating a strong sense of efficacy is set by Bandura (1994) as the experience of mastery. In other words, successful performance outcomes in the past can facilitate self-efficacy beliefs while poor performance outcomes hinder them. This source leads to the implication that learners' feeling of being a failure may come from their previous experiences and simple chances to replace their such memories with the new ones should be provided in the schools.

2.3.1.2. Vicarious Experiences: This source of efficacy stands for people's ability to learn from the outcomes of others' experiences. In other words, observing someone handling a task that is similar to the observer can help him or her feel that he or she is likely to perform the same task in the same way. Although weaker in potential compared to the mastery of experience, modelling or imitating is also effective on people's belief that they can competently perform in a given task. Stressing the significance of the similarity between the observer and the model, Bandura (1994) attracts the attention of the scholars to the impact of peers on students' behaviors in educational contexts.

2.3.1.3. Social Persuasion: Third source of self-efficacy beliefs, also termed as verbal persuasion, refers to direct encouragement or discouragement from another person to the individual's performance or ability to perform (Bandura, 1997). In this respect, those who are verbally persuaded to have the ability to be successful are more likely to sustain the task when even faced with social or personal difficulties. While discouragement is likely to decrease one's self-efficacy beliefs in being successful while the encouragement has the potential to do the opposite. This source brings about the pedagogical implication that assigns significance to the communication, with a special focus on careful use of verbal reinforcements, between the others and the students at school or home.

2.3.1.4. Physiological Arousal: The last source of self-efficacy beliefs stands for people's sensations from their body or emotional states (Bandura, 1994). In other words, people depend on their own bodies and how they perceive them to form their beliefs in their abilities to perform a desired task. That is to say, positive mood has a facilitating impact on self-efficacy while depressing mood has the tendency to lower it. As an example given by Bandura (1994), in activities requiring strength, people see their pain and fatigue as a sign of physical weaknesses resulting in poor self-efficacy. Mood also affects people's judgments of their personal efficacy. This leads itself to the pedagogical implication that students' stress must be reduced helping them feel at ease to perform academically better.

In conclusion, as stated by Bandura (1994), it is not the emotional or physical reactions that shape the self-efficacy belief but the way they are perceived and interpreted by the people. While those holding positive self-efficacy beliefs do not relate some affective factors to their abilities while those with lower self-efficacy factors interpret them as the main source of their inability to be successful in the given task.

As a closing remark, the four sources of self-efficacy beliefs, which are described above according to their strength of impact on self-efficacy respectively, are presented below.

Figure 2.2: Sources of Self-Efficacy

2.3.2. Significance of Self-Efficacy in EFL

The literature on learning regarding different fields of study has steadily revealed the significant relationship between psychological constructs and academic performance. Out of these constructs, self-efficacy appears as a factor closely related to goal setting and formation mediating "the effects of other potentially motivating variables, such as personality traits, feedback, participation in decision making, job autonomy, and monetary incentives" (Locke, Latham, 2006, 265). Additionally, it is set as an affective variables influencing and influenced by other similar variables ranging from motivation to anxiety and some personal traits such as gender, age and ability to succeed (Tilfarlioğlu, Cinkara, 2009).

Theoretically rooted in social cognitive theory of Bandura constructed in 1960s, self-efficacy has the potential to direct or determine cognitive, affective and behavioral processes operating in human bodies (Bandura, 1997). In this respect, Tilfarlioğlu and Cinkara (2009) exemplify this deep impact of self-efficacy beliefs on human processing as follows:

> "They affect whether individuals think in self-enhancing or self-debilitating ways; how well they motivate themselves and persevere when they face any difficulties; the quality of their emotional life and vulnerability to stress. Most individuals have knowledge and skills that are not used in proper settings. Therefore, the knowledge alone does not ensure effective practice" (30).

Compatible to the huge role it plays in various outputs of human beings as set by different scholars in the field, Bandura (1997) also touches upon the precious place of self-efficacy beliefs in learning process. In relation with other affective variables

such as attitudes, anxiety, motivation and cognitive competencies, self-efficacy comprises an important deal of the construct termed by Bandura (1997) as self-esteem. Having a great impact of how people perceive and handle with specific situations, self-efficacy acts as a consistent determinant of everything regarding learners' psychology, cognition and behavior during the learning process.

As in the case of overall learning, the relationship between EFL achievement and self-efficacy beliefs is also clear. Within these perspectives, in addition to wide range of variables from age, academic performance to learning strategies, Başaran and Cabaroğlu (2015) propose language as a variable that affects and is affected by self-efficacy beliefs of learners. Accordingly, research in EFL literature in the national and the international scope justifies the essential role of self-efficacy beliefs as one of the most significant predictor of second or foreign language learning achievement (Cotterall, 1999; Rahimi, Abedini, 2009; Tilfarlioğlu, Cinkara, 2009; Todaka, 2013). Consistent results acquired in the related literature posits that EFL learners with higher self-efficacy beliefs are more likely to be successful in language learning process than those with lower self-efficacy beliefs.

To conclude, a great deal of research conducted to explore the reasons behind differing degrees of success among EFL learners focuses on self-efficacy beliefs (Rahimi, Abedini, 2009). Revealed as one of the vital factor in academic performance in EFL, self-efficacy beliefs lead a number of scholars in the field to infer that learners' self-evaluations of their language ability to accomplish a desired task exerts decisive impact on their language learning performance.

2.3.3. Research on Self-Efficacy Beliefs in EFL

The field of English as a second language (ESL) and EFL embrace a large number of research studies focusing on affective factors such as self-esteem, motivation, self-efficacy, anxiety to delve into the variability among learners' levels of achievement (Tilfarlioğlu, Cinkara, 2009; Todaka, 2013). This body of research aimed to enlighten different aspects of the relationships among these factors comprising the total psychological construct. As a result, as proposed by Rahimi and Abedini (2009), important part of these studies in the literature revealed that affective variables play a more significant role in foreign or second language learning performance than some cognitive factors such as aptitude, intelligence and some

instructional factors such as teaching methods, techniques used and the time spent in or out of the classroom to attain the set goal. Among these affective factors, learners' beliefs, termed as self-efficacy, about their abilities to attain a desired goal in language learning process are regarded as one of the most essential contributor to language learning performance. Keeping its important role in EFL performance into consideration, some of the studies from national and international literature that dwell on self-efficacy in EFL were reviewed below to create a basis for comparing the related results of the present study.

Self-efficacy research has justified perceived self-efficacy belief as a remarkable variable in EFL (Başaran, Cabaroğlu, 2015). This widely accepted role of self-efficacy is clearly documented by the review of 32 studies between 2003 and 2012 by Raoofi et al (2013) to indicate how effective it was in ESL/EFL contexts. In this direction, the related literature touched upon many aspects of self-efficacy in EFL covering its relationship with learners' grade levels and their performances in addition to some other affective factors such as anxiety, attributions and motivation. Accordingly, Raoofi et al (2013) reported that 12 of 32 articles reviewed studied the role of EFL or ESL learners' self-efficacy beliefs in their performance in different language areas such as reading and listening.These studies, in harmony, revealed facilitating impact of positive self-efficacy beliefs on learners' language learning performance. On the other hand, 7 of the articles reviewed were explored to center on the relationship between either self-efficacy and anxiety or self-efficacy and personal attributions. The results of these studies shed lights on another aspect of self-efficacy indicating its significant negative relation to ESL or EFL performance. This is to say that learners holding positive beliefs in their abilities to succeed base their failure on their low level of effort while those with negative self-beliefs think they are unsuccessful not because of their effort but because of their lack of necessary ability.

Raoofi et al (2013) provided a summary of the review of studies conducted on self-efficacy in ESL or EFL in the international scope covering also those produced between 2003 and 2012 in Turkish context. In order to broaden the basis of comparison, the rest of the part dealt with research carried out after 2012. In this respect, another study led by Ghasembol and and Hashim (2013) provided an invaluable source. They enrich the basis on which the related results of the current

study will be compared. The reason behind this is the fact that the stated study was carried out with 187 teachers teaching English at language centres in one Middle-East country. In this respect, the study aimed to explore the relation between participant EFL teachers' sense of self-efficacy and their self-reported proficiency of English. With the adaptation of Teacher's Sense of Efficacy Scale and Teachers' reported English language proficiency questionnaire, the data was collected. As a result of correlation analyses, Ghasemboland and Hashim (2013) explored that the participant teachers' perceived self-efficacy were significantly related to their self-reported proficiency in English.

Another study led by Doordinejad and Afshar's study (2014) conducted with the participation of 400 Iranian high school students adds to the basis on which the related results of the current study will be compared. In their study, the researchers (2014) attempted to investigate the relationship between self-efficacy and EFL performance among third grade high school students. Foreign Language Learners' Self-efficacy scale for Iranian Students was employed to assess the participants' self-efficacy beliefs while the third-grade English language final exam prepared by Iran Ministry of Education was used to document their performance levels. They, as a result, explored that self-efficacy beliefs are significantly related to EFL high school students' performance.

A very recent study carried out by Osman et al (2016) detailed the place of self-efficacy in EFL. In their comprehensive study, the researchers (2016) examined the reciprocal relationships among gender, EFL reading and self-efficacy beliefs in EFL. In this study, 260 grade four students and 376 grade ten students were asked to complete reading achievement test developed by the Ministry of Education and a reading self-efficacy beliefs scale developed by the researchers whose reliability and validity procedures were realized. Within the interest of the current study, the results indicated that there was a strong relationship between self-efficacy in EFL reading and EFL reading performance in both four and ten grade students. In other words, students with higher levels of self-efficacy beliefs in both grades were found to significantly better EFL readers than those with lower levels of self-efficacy. As a result of this study, the consistent impact of self-efficacy on EFL performance was justified with a different age group in a different language skill in the favour of females.

Thanks to the review of some studies on self-efficacy in EFL in international context, it was consistently asserted that learners' beliefs in their abilities to succeed are significantly related to their performance to succeed. However, the number of similar studies in Turkish context is limited (Tilfarlioğlu, Cinkara, 2009). With the aim of forming a well-qualified basis of comparison with the results generated by the present study, some representatives of these few studies on self-efficacy in EFL in Turkey were briefly dealt with below.

One of the studies, as covered in the review of Raoofi et al (2013), was led by Tilfarlioğlu and Cinkara (2009). As one of the main representative studies that well draw the boundaries of self-efficacy research in EFL in Turkey, Tilfarlioğlu and Cinkara (2009) surveyed 175 EFL students from Gaziantep University, Turkey. The main aim of the study was to conceptualize the relation between EFL students' self-efficacy beliefs and their level of attainment in EFL learning process. The related data was collected through the use of a 7 Likert Self-Efficacy Scale and the end of year GPA scores. In a similar vein to those results produced by international studies, correlation analyses set the significant relation between reported self-efficacy beliefs in EFL and level of achievement in EFL.

A very recent self-efficacy research in Turkey aimed to shed light on its relation to another important psychological factor, anxiety (Meriç, 2015). In this research, the researcher attempted to explore the relation between EFL student teachers' levels of anxiety and their self-efficacy beliefs in teaching. 117 students from Anadolu University participated in the study and they completed a *Foreign Language Student Teacher Anxiety Scale* (FLSTAS) and a *Self-Efficacy Questionnaire*. As a result of the administration of these instruments in addition to semi-structured interviews, no gender impact on neither anxiety nor self-efficacy was found. On the other hand, as more related to the focus of the current study, anxiety and self-efficacy were found to be correlated. In this respect, this result seems to be in parallel with Bandura's (1997) assertion that self-efficacy is influential in the amount of anxiety one feels while handling with a task.

Another recent study was carried out by Tuncer and Doğan (2016) broadened the scope of study drawn by Meriç (2015). In their study, Tuncer and Doğan (2016) attempted to shed lights on the relation among self-efficacy, anxiety and metacognitive awareness of 271 EFL students at Fırat University, Elazığ in Turkey.

Through the use of the *Foreign Language Classroom Anxiety Questionnaire*, *the Academic Self-efficacy Questionnaire and the Metacognitive Awareness Inventory,* the data was collected. As a result, structural equation model, similar to Meriç (2015), indicated that self-efficacy predicted both EFL anxiety and metacognitive awareness.

In conclusion, this brief review of research on self-efficacy in EFL has clarified the picture regarding the role of self-efficacy in many facets of EFL ranging from its interaction with other affective variables to such cognitive factors as metacognitive awareness. The results of this bulk of studies have consistently revealed self-efficacy in combination with the outsets of motivation as a strong predictor EFL achievement. In the similar way, self-efficacy has a vital role in EFL learners' performance in Turkish contexts. This role can be attributed to their constant fears of failure arising from a number of central exams they have to take during their educational lives. In contrast to this fear, studies strengthened its impact on EFL performance by explaining much of that performance with not competence but self-efficacy. Drawing on these consistent results on self-efficacy in EFL, the scholars are motivated to make some pedagogical implications. Regarding this, it is highly implied that it is significant to design learning environments and activities in the classroom in a way to enhance students' self-efficacy beliefs that would lead itself to motivation and ultimately achievement in given EFL tasks.

To sum up, the present study takes the strongly cited enhancing power of self-efficacy on EFL performance into account. Within this respect, the researcher dwells on it in combination with attitudes toward EFL as a key of solving failure in EFL in Turkish context. With this purpose in mind, in the current study, flipped classroom model, adopting an eclectic approach, is associated with self-efficacy and attitudes as vigorous predictors of EFL success.

To conclude, dwelling on their high potentials, the present study takes the form of the first attempt to embrace self-efficacy beliefs, attitudes and flipped classroom in the same pot. Accepting not only *self-efficacy and attitudes* as vigorous predictors of EFL success but also education as a long process, the present study hypothesizes that flipped classroom model will initiate learners to devote more of their time to learn in and out of classroom. This will, in a cyclical manner, help them form positive attitudes toward and stronger beliefs in their abilities to succeed in EFL in the long

run. By adopting a mixed method design, the researcher, compatible to the dynamics of 21st century, proposes the combination of these variables as an efficient way of dealing with failure in EFL in Turkey by better responding the needs of students as digital natives. In essence of these, the current mixed method study takes the form of the pioneering attempt in Turkey to enlighten the concerned bodies ranging from parents, students, teachers to curriculum developers and policy makers about the relationship among flipped classroom model, EFL performance, self-efficacy and attitudes as the consistently cited determinants of this performance.

3. METHODOLOGY

The main aim of this study is to explore the impact of flipped classroom model on EFL in Turkey. Within this perspective, the related literature was explored to sign negative attitudes toward, lack of self-efficacy beliefs in it, necessary time and technology as the predictors of the failure in EFL in Turkey. Based on this reality, the researcher sought to enlighten the concerned bodies about the impact of flipped classroom model on those variables as the facilitator of EFL performance.

To serve for the stated aim of the study, it was necessary to collect data that indicates the participants' EFL performance, their attitudes toward and self-efficacy beliefs in EFL. In this respect, the data collected through the application of *EFL Achievement Test* developed by the researcher, *Self-Efficacy Belief in English Scale* (Yanar, Bümen, 2012), *Attitudes toward English Scale* (Aiken, 1979; Tunç, 2003; Tuncer, Berkant, Doğan; 2015) as pre- and post-tests stood for the quantitative data of the study. On the other hand, as the facilitator of the data, qualitative data attained through semi-structured interviews developed by the researcher was collected to indicate the learners' perceptions of their flipped classroom experience.

With this feature of the study, the researcher promises valuable information by combining the recent teaching model compatible with technology, the flipped classroom teaching model and a bleeding wound in of the country, EFL in Turkey.

Within this respect, this part of the study will draw the framework of the study ranging from the design, the participants, the pilot study data collection instruments, procedures to data analysis.

3.1. Design

The main purpose of the study is to explore the impact of flipped classroom teaching model on the university prep EFL learners' academic performance and their attitudes toward and self-efficacy beliefs in EFL. In this respect, this embedded study will be grounded in a mixed method design. Inspired by the assumption that the uses of both

qualitative and quantitative data will help "better understand the research problem and question than either method by itself", the mixed method research design is based on blending of qualitative and quantitative methods in a single study (Creswell, 2012, 535). In order to develop a more in-depth understanding of the experimental outcomes of the study, the embedded design out of mixed method research designs was adopted. Creswell defines embedded studies as a form of mixed method design where "the purpose is to collect quantitative or qualitative data simultaneously or sequentially, but to have one of data play a supportive role to the other form of data" (Creswell, 2012, 544). Creswell (2012) explains that the pilot supportive data can be in any form while the commonly seen trend in the literature is to use qualitative data as the facilitator of the quantitative data. As an example, he (2012) adds this type of design can be drawn when "the quantitative data will address whether the intervention had an impact on the outcomes, whereas the qualitative data will assess how the participants experienced the intervention" and "explain and follow up on the quantitative outcome results" (544-545).

In accordance with theoretical explanation about the model, the current study was conducted in two phases. While the first phase was grounded in quasi-experimental pre- and post-test design, the second phase will be conducted in phenomenological design. The first experimental design aimed to explore the impact of the intervention on experiment-group EFL learners' academic performance, attitudes toward and self-efficacy beliefs in EFL. Quasi-experimental design was held consciously instead of true experimental design in the first fold since it does not include "random assignment of participants to groups" (Creswell, 2012, 309). As set by Creswell (2012), in educational surroundings such as schools, colleges or universities, educators and researchers often use intact groups in experiments since randomly assigning them to experimental and control groups may disrupt classroom learning. Dwelling on this, the researcher included already existing two groups of EFL learners in the study. Yet, to deal with threats posed by this design, the researcher ensured that participant two groups of learners were equal in terms of the concerned variables of the study.

On the other hand, the second phase served as facilitator of the quantitative outcomes in the form of follow-up semi-controlled interviews with the experiment group students about their flipping experiences. Therefore, the second phase of the study

was grounded in phenomenological design since it aims to have deep information about how people experience, value and perceive a certain phenomenon through some qualitative methods such as interviews and observations (Hycner, 1985). As a result of this interest in the individual experience, phenomenology produces outcomes about people's motivations and actions and determines conceptual understanding of a certain group of participants about a specific experience (Çekmez, Yıldız, Bütüner, 2012)

Within this perspective, the flow of the current mixed design of the study is illustrated as follows:

Figure 3.1: Flow of the Design

Serving for the internal validity of the design, the courses in both groups were led by the researcher himself. In this respect, the students in the experimental group were taught based on the Flipped Classroom Model during the research process in the fall term of 2016-2017 academic year, while those in the control group were taught following traditional lecture-based instruction. From the first week of the research process, lecture parts of the lessons in the experimental group were flipped with the related homework sections. The students were asked to watch the video lectures created by the researcher or provided by the researcher from different sources. These were shared with the students through Edmodo, a network platform where teachers can collaborate with their students and check their assignments and developments. On the other hand, the students in the control group were taught traditionally based on lecturing from the first week of the semester. They were asked to complete their homework at home. However, the same syllabus was aimed to be covered by the researcher in both groups (see Appendix 16). .

To conclude, in a route from quantitative to qualitative data, the researcher dwells on the strengths of both types of data to promise deep and valuable data to wide range of concerned bodies.

3.2. Participants

In experimental designs, it is often impossible to randomly select the participants for a number of reasons. Therefore, the participants are selected from those who are both available and willing to take part in the experiment (Creswell, 2012). Within this perspective, the researcher employed *convenience sampling* method to draw the sample of the study. This is usually adopted when the researcher aims to select individuals who "are available, convenient and represent some characteristics the investigator seeks to study" (Creswell, 2012, 145). Such a sample "can provide useful information for answering questions and hypotheses" (Creswell, 2012, 145). As a result of this technique and such constraints as time, fund and nature of the data collection instruments, the researcher included prep EFL students from two different classes at Gebze Technical University, whose populations range from 20 to 34. The ages of the participant students range from eighteen to twenty. Among the available classes, two ones where the researcher teaches himself were adopted. They were exposed to pre-test administrations of EFL Achievement Test, Attitudes toward EFL and Self-Efficacy Beliefs in EFL scales. To see if these classes are equated in terms of the variables within the interest of the study, the researcher computed necessary analyses through *SPSS 21 for Windows*. In this respect, it was necessary to decide on the type of the data since each one of them require different type of analyses. Therefore, the researcher relied on normality test as one of the accepted parameters that help to decide on the type of the data. Hence, assumptions that assess the normality of the distribution of the data gathered in the pre-test application of the instruments were covered.

After finding that the data distributed normally, parametric tests were determined to be employed in order to seek for any difference between two groups of students in terms of the concerned variables of the study. Relying on this, the groups were equated through parametric tests.

Moreover, these EFL learners of the study were informed about the purpose of the study and requested to sign Informed Consent Form summarizing their rights and the

framework of the study itself before they were involved in the study (see Appendix 13). The experimental procedures of the study were designed to replace those of the course activities the students have to participate. Therefore, the researcher asked the students to sign the form to show their consent not to involve in the study but to permit the researcher use their data in the research. Yet, it was ensured that all individual opinions, performances were confidential and it was their right not to participate in the follow-up interviews of the study.

Eventually, the distribution of the participants according to the equated control and experimental groups was shown in the following table.

Table 3.1: Descriptive Statistics for the Participants

Groups	Gender	N	%	$\%_{valid}$	$\%_{cumulative}$
Experiment	**Male**	12	29,3	29,3	29,3
	Female	9	21,9	21,9	51,2
Control	**Male**	11	26,9	26,9	78,1
	Female	9	21,9	21,9	100

As seen in the table, 12 of the students in the experimental group were male while the rest (9) were females. On the other hand, 26, 9 % (11) of the participant students were males in the control group while the remaining 9 students were females.

3.3. Data Collection Instruments

The purpose of this study is to shed lights on the impact of flipped classroom on attitudes toward, self-efficacy beliefs in EFL covering the advantages, perceptions and academic outcomes of the flipped classroom teaching model. To realize this aim of the research, it was a must to obtain the participant EFL learners' pre- and post-test scores of academic EFL performance, attitudes toward and self-efficacy beliefs in EFL in addition to their perceptions of the flipped classroom model. In other words, the data collection procedures of this study were two folded. Within this frame, the quantitative data was collected by the administration of *EFL Achievement Test* developed by the researcher, *Self-Efficacy Belief in English Scale* (Yanar, Bümen,

2012), *Attitudes toward English Scale* (Aiken, 1979; Tunç, 2003; Tuncer, Berkant, Doğan; 2015) as pre- and post-tests in the fall term of 2016-2017 Academic Year. On the other, as the facilitator of the data, qualitative data that indicate the learners' perceptions of their flipped classroom experience was attained through semi-structured interviews with nine students in the experiment group. In this perspective, the following section is allocated for the comprehensive explanation on these data collection instruments.

3.3.1. EFL Achievement Test

In order to elicit the participant EFL learners' scores of EFL performance, the EFL Achievement Test developed by the researcher was applied (see Appendix 1). This test aimed to explore the progress the participants make during the English course in a prep EFL class at Gebze Technical University. Therefore, this test is expected to be sensitive enough to gain diagnostic information about students' progress in EFL during a term. This makes it obligatory to obtain the results as detailed and accurate as possible.

Based on the purposes stated above, the results of the test are expected to help teachers make important decisions about their students and instruction. Therefore, the researcher must be sure that the test measures the adequate sampling of classroom content and the cognitive processes implemented during the instruction (Fives, Barnes, 2013). These all serve to improve the "validity of teachers' evaluations or decisions about their students based on the given assessment" (Fives, Barnes, 2013, 3). However, there is no precise set of steps accepted by experts to follow during the process of constructing an achievement test. Yet, to ensure the quality of the test to be developed, the researcher of the present thesis followed the stages in the form of the composition of Turgut, Baykul (2012) and Ivanova (2011). Respectively, the process of constructing the achievement test followed the stages of stating the overall plan of the test, content definition, test specifications, item writing, test design, item reduction, piloting the test, scoring test responses, item analysis and forming the ultimate test. These stages and the procedures realized within those stages were covered below:

3.3.1.1. Overall Plan: The first step of developing a test is setting its goal so that it can overlap with its purpose serving for high validity. In this respect, university prep EFL achievement test was aimed to measure EFL students' progress during the fall term at Gebze Technical University.

3.3.1.2. Content Definition: Ivanavo (2011) indicates the second step as a stage where the test developer needs to determine the content to be tested. Keeping this in mind, the objectives and functions specified in the English Language Curriculum of the prep classes at university were reviewed. Then, in addition to the syllabus, the coursebooks and the objectives that are to be covered during the term were reviewed. Following this, the most critical objectives were identified. To ensure the match between the objectives and the items that test the acquisition of those objectives by the test takers, at this stage, a table of specifications was developed by the researcher, the test developer (see Appendix 4). In this way, the test developer, the researcher himself, aimed to bridge instruction, assessment and support teachers' judgments about their students and the quality of the instruction itself.

3.3.1.3. Test Specifications: Based on the purpose, content of the test demonstrated through table of specifications, test developers are advised to make a detailed plan of the test ranging from the addressee to the skills stressed in each part of the test. In this plan, to see how much of these objectives are attained by the students, the number of the test items to include were made clear by the researcher (see Appendix 3). While making those decisions with regard to the distribution of items and their types and numbers, such points as the number of objectives, teacher's understanding of the students, the class time allocated for testing were taken into account. Embracing these points, a balanced skills-based approach (listening, grammar, reading, vocabulary and writing) was adopted. As a result, each of the skill constituted 20 % of the whole scoring.

3.3.1.4. Item Writing: At this stage of test development, the test developers are advised to take the test specifications and the table of specifications into account to come up with initial set of test items (Turgut, Baykul, 2012). Within this perspective, the researcher developed a skill based mixed type test in a way that would cover each of the objectives in the specification of the test. Taking the reliability procedures into account, which may force the developer to remove some of the items due to their

94

poor quality, in the initial procedure for developing the achievement test, the researcher aimed to include at least items two times more than planned items to test each of the objectives in the table of specifications. As a result, the researcher included test items ranging from 2 to 8 to test each of the critical objectives.

3.3.1.5. Test Design: At this stage, the test is put into its final form to pilot. All the test items are ordered properly in a way that will not distract the test takers. Therefore, following Turgut, Baykul (2012), the researcher asked two colleagues to check the test for any grammatical or technical mistakes. Taking their suggestions, the researcher put the finishing touches on the test "to maximize the ease of reading and thus to minimize the time necessary to complete the items" (Ivanova, 2011, 283).

3.3.1.6. Item Reduction: In addition to edition on the formatting of the test, the researcher also needs to worry about the validity of the test to be developed from the perspective of the experts. In this respect, it must be ensured if the content of the test corresponds to content specified in the table (Ivanova, 2011). To set the content validity of the test, two folded procedures were realized collecting data from five EFL teachers and five experts in EFL. Following adaptation of Lawshe technique, in the first fold, five EFL teachers from the university were asked to evaluate the content validity of the objectives found critical by the researcher among those to be covered during the term. They were asked to comment through *Evaluation of Functions Forms* where they were asked to assign numbers to their opinions from 1 to 5 for each of the functions specified by the researcher (see Appendix 5). In the second fold, to test whether the objectives in the table of specifications match with the items planned to test them and the levels they were specified to correspond at Bloom's taxonomy, five experts in EFL from different universities were asked to evaluate the harmony between the designed test and the table of specifications. Instead of limiting them to three classical options as necessary**,** unnecessary and useful but not necessary, on an *Evaluation of Achievement Test Form*, they were asked to mark from 1 to 5 to give a statistical value for the content and face validity of the test (see Appendix 6). While 1 was standing for "completely invalid", 5 was indicating "completely valid" in the form. They were also asked to comment on anything else that they find important to edit on the test or the table of specifications. Then, simple descriptive calculations in the form of mean and percentages were

followed. In evaluating these results, the number of gaps and options were used as criteria (Tekin, 1996). In this respect, the following ranges of mean scores were taken as a basis:

1.00-1.80 Completely invalid

1.81-2.60 Invalid

2,61-3.40 Undecided

3.41-4.20 Valid

4.21-5.00 Completely valid

As a result, the test looked like a valid measure of EFL performance of prep EFL learners at Gebze Technical university regarding the objectives specified with the mean score of 4,8 out of 5 according to 5 EFL teachers from the university. This indicated that none of the objective specified as critical by the researcher was found to be invalid by 5 EFL teachers. On the other hand, the experts' evaluations of the harmony between the items, functions and their corresponding stages at Bloom's taxonomy produced 4,3 out of 5 on content validity and 4,6 on face validity. In other words, this meant that the objectives specified by the researcher among those to be covered during the term were justified by EFL teachers at the university. Yet, their comments were taken into account to edit the objectives further. On the other hand, the test items developed to measure the acquisitions of those objectives and their correspondences at the taxonomy were also confirmed by expert views. As a result, thanks to the specification of the test and opinions provided by the experts and the teachers, the necessary changes were made and the content validity of the test was ensured.

3.3.1.7. Pilot Testing: This stage includes the production, publication and administration of the test. Within this perspective, the researcher took the suggestions provided by the teachers and the experts to finalize the trial form. In this respect, the test was set to be administered in two sessions. The first session was decided to last for 60 minutes covering listening, grammar, reading, vocabulary while the second session was determined to cover writing in 30 minutes. However, some items found by the teachers difficult for the students were not removed from the test before the pilot study. The reason behind this was that pilot study was a small scale study that

may enlighten the researchers about such facets of the real study as feasibility, cost, time, reliability and validity of the data collection instruments (Van Teijlingen, Hundley, 2002). In this respect, in order to be aware of the potential problems that may come out during the implementation of the study, the researcher conducted pilot study with 38 students who were qualified to study at English Prep Schools of eight different universities (5 state and 3 private) in Turkey in 2016-2017. As a result of this, the researcher did not only come up with a set of data to calculate item analysis but also found that the time limit set for the testers to complete the test was enough.

3.3.1.8. Scoring Test Responses: This stage of the test construction process stands for ensuring the accuracy of testing scores with a detailed scoring key (Ivanova, 2011). Based on this, the researcher provided a detailed answer key (see Appendix 2). The scoring key was prepared as detailed as possible in order to attain objective scoring. To evaluate students' levels of success in each skill, nothing was deducted for the errors of the skill outside the confines of what the item tested.

Due to its open ended and creative nature, the writing part was assessed by three EFL teachers following the analytic rubric provided in the detailed answer key. In the rubric, four areas (grammar, vocabulary, content and organization) were taken into consideration. Each main area was assessed following the guidelines where each performance in each of four criteria was analytically and separately evaluated out of 5. The final writing performance of every test taker was produced capturing the mean of three EFL assessors.

3.3.1.9. Item Analysis: This stage of the test development stands for dwelling on the set of data acquired from the pilot study. At this stage, the test developers analyze the performances on each item to produce some indices that may help them evaluate the overall quality of the test. Such indices as item mean score, item standard deviation, the item difficulty and discrimination guide the test developer to decide on the items that need to kept, removed or edited (Turgut, Baykul, 2012).

In this study, the reliability analysis of the achievement test, including item difficulty and discrimination were run by *Item and Test Analysis Program–ITEMAN*. ITEMAN was adopted in the study since it analyzes test providing conventional item analysis for each item through such statistics as item-total correlations, discrimination and difficulty index and the test as a whole through mean, median,

standard deviation and the reliability. Since ITEMAN allows the researchers to analyze the quality of dichotomously scored items, the researcher of the present study coded the data dichotomously as correct and incorrect. While 1 was used for correct responses, 0 was selected to code incorrect responses. ITEMAN was also expected to help the researcher clearly indicate the impact of each item and even response alternatives on reliability of the test by providing the researcher with KR-20 reliability (Thompson, Guyer, 2010). This was realized by the researcher consciously since KR-20 is stated to be a measure of internal consistency for tests or surveys which are composed of dichotomously scored items (Bademci, 2011; Tan, 2009). In this way, the researcher utilized ITEMAN rooted in conventional test theory to include appropriate items in the final draft of the test based on their difficulty and discrimination power.

3.3.1.10. Forming the Ultimate Test: The last step of constructing an achievement test is to form the ultimate test by including the items having difficulty (P) and discrimination (Rpris) indices within acceptable ranges.

Item difficulty index stands for the percentage of the students who answered the item correctly to the number of all the test takers (Turgut, Baykul, 2012). This item difficulty index ranges from 0.0 to 1.0. While 0 shows there is no student answering the item correctly, 1.0 indicates there is no student answering the item incorrectly. On the other hand, item discrimination index (Rpbis) indicates the power of an item to measure what it aims to measure by discriminating between students with higher scores and those with lower scores. Ranging from (-) 1.0 to (+) 1, the index shows the correlation of the item with the whole of the test. While (-) values is a signature of mismatch between the item and the rest of the test, (+) values indicate a harmony of the item with the whole test. On the other hand, a discrimination index at 0 shows there is no relation between what the item and the rest of the test measures.

Based on the framework provided above, the achievement test was finalized by taking results of item analysis. While accepting or rejecting the items for the ultimate test, the acceptable discrimination and difficulty ranges put forward by Turgut and Baykul (2012) as a composition of different views in the literature were adopted. In this respect, for item discrimination index (Rpbis), if the discrimination value is;

- 0,19 and lower, the item must be discarded from the test,

- between 0,20 and 0,29, the item must be edited or revised,
- 0,30 and over, the item is acceptable.

In terms of item difficulty index, if the difficulty index is between;

- 0,00 – 0,19, the item is very difficult,
- 0,20 – 0,34, the item is difficult,
- 0,35 – 0,64, the item is moderately difficult,
- 0,65 – 0,79, the item is easy,
- 0,80 – 1,00, the item is really easy.

Taking these difficulty ranges and Turgut and Baykul's (2012) remark on item difficulty proportion of test items into account, 70 % of the test items was aimed to be moderately difficult, while 13 % of them was supposed to be difficult. 13 % of the test items was targeted to be easy while 2 % of the items was aimed to be very difficult. The remaining 2 % was planned to be very easy. In this way, it was attempted to address all types of test takers.

As a result of following the aforementioned steps of test construction including issues related to the validity and reliability procedures administered within the present study, the researcher developed and ensured *EFL Achievement Test* as a valid and reliable assessor of EFL prep learners' EFL performance at Gebze University ($KR_{20}= 0,91$).

3.3.2. EFL Self-Efficacy Belief Scale

To find out the impact of flipped classroom teaching model on EFL learners' self-reported efficacy beliefs in EFL, it was necessary to employ a self-efficacy scale as a pre- and post-test. Hence, *Self-Efficacy English Scale* developed by Yanar and Bümen (2012) was administered to 38 prep EFL learners at Gebze Technical University (see Appendix 7). This scale originally aimed to measure high school EFL students' reported self-efficacy beliefs in EFL ($\alpha = 0.97$). More specifically, this scale aimed to explore EFL learners' self-reported efficacy beliefs in four main ability domains of English as listening, reading, speaking and writing. The original scale consisted of 34 items which are scored on a 5-point Likert-type scale ranging from "1= strongly disagree" to "5= strongly agree". To test the content validity of the scale, the researchers consulted 14 experts including two from the assessment and

evaluation department, three from curriculum and instruction and one from counselling and guidance departments. The reliability statistics of the revised form of scale was held using a number of Anatolian High School EFL learners at eleventh grades (n=296). The results of explanatory factor analysis explored that 34 items with extraction values ranging 0.42 to 0.69, as a result of omitting 13 items having low factor loadings, included in the scale gathered in four factors. Following this, the researchers also employed confirmatory factor analysis to confirm the four-sub-scaled model. As a result of the fit indices, the researchers confirmed that self-efficacy beliefs of the EFL learners gathered around four areas as set by the explored model (RMSEA=0.044 and SRMR=0.046). The internal consistency reliability coefficients for each of four subscales were also found to be within acceptable ranges as follows: reading ($\alpha = 0.92$), writing ($\alpha = 0.88$), listening ($\alpha = 0.93$) and speaking ($\alpha = 0.92$).

Keeping the impact of self-efficacy beliefs on achieving the purpose in EFL in mind, the researcher adopted the aforementioned scale developed by Yanar and Bümen (2012). In addition to the elaborate validity and reliability procedures followed by the researchers, one of the pioneering reason behind this choice is that this scale was developed for high school students who are around the ages of those at university displaying similar physical and psychological attributes. In addition to addressing the age groups of the participants of the present study, as noted by Yanar and Bümen (2012), other self-efficacy scales were either other subject areas specific or derived from motivational scales. However, the adopted scale is the first one focusing on four main sub-skills in EFL (Yanar, Bümen, 2012). When it is noted that the current study seeks to explore the impact of flipped classroom model on four skills in EFL, skill-based aspect of the scale also makes a perfect match with the purpose of the study.

In the current study, *Self-Efficacy English Scale* (Yanar, Bümen, 2012) was tested for its statistical relevance to the sample of the study. Within this respect, the first step to follow is to get the permission of the test developers by informing them where and how to use the scale (Karakoç, Dönmez, 2014). Therefore, the researcher obtained the permission from Bümen to administer the scale in the current study (see Appendix 8). Then after, the researcher initiated the analysis of the data gathered from the pilot study carried out with 38 students who were qualified to study at English Prep Schools of eight different universities (5 state and 3 private) in Turkey

in 2016-2017. Before checking the reliability and validity of the instrument, as a prerequisite, first of all KMO & Bartlett's Test of Sphericity was computed to measure the adequacy of the sampling (KMO = 0,55; x^2 = 1214,138; p < .001).

Once the sample was found to be suitable to conduct intended analyses, the researcher set to test the validity of the modified scale. Golafshani (2003) refers to the validity in quantitative research as "construct validity". The construct of the scale refers to "self-efficacy beliefs in EFL". In this respect, the researchers administered confirmatory factor analyses (CFA) to produce sound outcomes if the adopted scale is a valid measure of the stated construct within the sampling of the present study. Yanar and Bümen (2012) computed exploratory factor analysis (EFA) in the phase of developing a scale for the first time to explore the possible underlying factor structure of a set of observed variables without a preconceived structure on the outcome (Child, 1990). In this respect, in the present study, the researcher used *LISREL 8.54 for Windows* and more specifically a full structural equation model (SEM) to test and confirm the four-dimensioned model of the adopted scale with the participants similar to those of the current study in the pilot study. This is because CFA is the sum of a number of statistical tests to determine the adequacy of the model and confirm the set relationship between observed variables and their underlying latent constructs (Bryant, Yarnold, 1995). In other words, CFA is employed to confirm theoretical structures explored as a result of EFA (Bektas, Akman, 2013). As a result of these, CFA was administered to the factorial structure explored before to explain 61,41 % of the total variance.

Following maximum likelihood estimation method, in addition to path diagrams, fit indices were taken into account to evaluate the results related to the model. Because fit indices serve to explore if a CFA model fits with the data or the variables (Bryant, Yarnold, 1995). Out of these fit indices, the following values were analyzed in order to make a conclusive remark on the scale: Chi-square (x^2), RMSEA (Root Mean Square Error of Approximation), SRMR (Standardized Root Mean Square), RMR (Standardized Root Mean Square), CFI (Comparative Fit Index), GFI (Goodness of Fit Index), Normed Fit Index (NFI), Incremental fit index (IFI) and AGFI (Adjusted Goodness of Fit Index).While deciding on the fit of the explored model, the following ranges specified in the table were used.

Table 3.2: Acceptable Fit Indices for CFA

Fit Indices	Acceptable Values
Chi-square (x^2) / df	≤ 5
GFI	≥ 0.90
AGFI	≥ 0.85
CFI	≥ 0.90
RMSEA	$\leq 0,08$
SRMR	$\leq 0,08$
RMR	$\leq 0,08$
NFI	≥ 0.90
IFI	≥ 0.90

Adapted from Harrington, Donna. 2008.**Confirmatory Factor Analysis**. USA: Oxford University Press; Schumacker, Randall, E.,Lomax, Richard, G. 2004.**A Beginner's Guide to Structural Equation Modeling**. Second edition. Mahwah, NJ: Lawrence Erlbaum Associates; Şimşek, Ömer, Faruk. 2007. **Yapısal Eşitlik Modellemesine Giriş: Temel İlkeler ve LISREL Uygulamaları**. Ankara: Ekinoks Yayıncılık

In addition to CFA analysis run to confirm the explored model, the next step was to compute the reliability coefficient of the scale. Therefore, cronbach's alpha (α) value using SPSS *version 21 for Windows* was computed to test the reliability of the scale (α = 0.94).The reason behind computing cronbach's alpha instead of KR_{20} is that alpha is the most common measure of internal consistency especially with multiple likert questions in a survey/questionnaire that form a scale and when the aim of the researchers is to determine if the scale is reliable (Bland, Altman, 1997; Cronbach, 1951).

On the other hand, in the present study, the students' scores of self-efficacy beliefs were interpreted following Yanar and Bümen (2012), who suggested three types of self-efficacy beliefs in EFL: high (mean of 4.1 or higher), medium (mean of 3 – 4), and low (2.9 or lower).

To sum up, employing all the procedures related to getting permission, CFA and cronbach alpha coefficients analyses confirmed that the adopted scale of self-efficacy including 34 items in four factors as a valid and reliable assessor of university prep EFL learners' self-reported efficacy beliefs in four main skills of English (RMSEA=0.080; SRMR=0.013; $\chi2$ / df = 1,24; α = 0.94).

3.3.3. Attitudes toward English Scale

In addition to self-efficacy beliefs, attitudes can predict the success of the learner in EFL (İlhan, Karataş, 2015). In this respect, in order to see the impact of the flipped classroom treatment on the participant EFL students' attitudes toward English, the scale rooted in 'Attitudes toward Mathematics and Physics Scale' developed by Aiken (1979). The original scale consisted of 24 items scored on 5-point likert type. As a result of explanatory factor analysis, Aiken (1979) explored that 24 items gathered around four factors: Love for Math, Fear of Math, Motivation for Math and Significance of Math. Under every factor, there were six items, three of which were negative while the remaining three were positive. Aiken (1979) found the instrument was a reliable assessor of Iranian EFL learners' attitudes toward EFL (α = .81).

Tunç (2003) acquired Turkish adaptation of this aforementioned scale in a way to determine university prep EFL learners' attitudes toward English. By being loyal to its original form, Tunç (2003) modified the original four factored scale to match up with EFL in one factor under 24 items to ensure that the students can easily circle the number in the scale that they think best reflects them. In this respect, the scale went through back and forth translation processes. Tunç (2003), therefore, included such general statements regarding the EFL as 'English is not a very interesting class' and 'I like doing English exercises'. Just as the original scale, the adopted scale consisted of 24 items which are scored on a 5-point Likert-type scale ranging from "1 = strongly disagree" to "5 = strongly agree". (Tunç, 2003). As a result of the internal consistency reliability coefficient analysis, cronbach's alpha value (α= .77) was found to be within acceptable range (Tunç, 2003). Yet, the scale was not validated by Tunç and it was found to be utilized as one-factor scale. In contrast to lack of validating results, the reliability of the scale was also confirmed by some other studies conducted by Çelebi (2009) and Çimen respectively (2011) (α= .87; .89)

Elaborate validity and reliability procedures of the scale were conducted by Tuncer, Berkant, Doğan (2015). The researchers employed both explanatory and confirmatory factor analyses on the data collected from 271 university prep EFL learners. Compatible to its original form developed by Aiken (1979), the results of explanatory factor analysis indicated that the scale had four factors with 19 items explaining the 54.180 % of the total variance (α= .88). In order to test this model, the researchers also computed confirmatory factor analysis statistics and confirmed the four-factored and 19-item structured model of the instrument as a valid and reliable assessor of university prep EFL learners' attitudes toward English (RMSEA=0.059; SRMR=0.061; α =.865).

In the current study, Attitudes toward English Scale was administered as pre- and post-tests to gain insights about the impact of the flipped classroom on the learners' attitudes toward EFL (see Appendix 9). This scale was adopted not only because it was deeply rooted in one of the pioneering scale in the history but it was also set as reliable in many studies conducted with similar groups of participants to those of the current study (Çelebi, 2009; Çimen, 2011; Tunç, 2003). Moreover, it also went through recent detailed factor analyses with the participation of university prep EFL learners (Tuncer, Berkant, Doğan, 2015). However, to produce unbiased results, the scale was again tested for its statistical relevance to the sample of the study. Hence, same procedures followed to test the relevance of the Self-Efficacy English Scale were repeated in the same order. Therefore, first, the researcher aimed to obtain the permission to administer the scale in the current study from the three developers of the scale through e-mail. Berkant and Tuncer answered the mail and allowed the researcher to use the scale being loyal to ethical issues related to citations (see Appendix 10).Then, in the pilot study carried out with 38 students registered at English Prep Schools of eight different universities (5 state and 3 private) in Turkey in 2016-2017, KMO & Bartlett's Test of Sphericity was computed to measure the adequacy of the sampling (KMO = 0,732; x^2 = 434,477; p < .001).

After setting the adequacy of the sampling, the researcher stepped forward to compute confirmatory factor analyses to strengthen the comment if the adopted scale is a valid measure of the "attitudes toward EFL" construct within sampling of the current study. Since EFA was run by Tuncer, Berkant, Doğan (2015) in the developmental period of the scale, the researcher used *LISREL 8.54 for Windows* and

SEM to test and confirm the four-dimensioned model of the adopted scale with the participants similar to those of the current study in the pilot study.

Following maximum likelihood estimation method, fit indices specified in table 3.4were taken into account to evaluate the results related to the model.

In addition to CFA analysis run to confirm the explored model with four factors around 19 items, the next step was to compute cronbach's alpha (α) value of the scale using SPSS version 21 for Windows (α = 0.91). The reason behind computing cronbach's alpha was the same as that of the EFL Self-Efficacy Belief Scale. Internal consistency of the scales with multiple likert questions is determined relying oncronbach's alpha (α) value.

On the other hand, in the present study, the students' scores of attitudes toward English Scale were interpreted following the formula based on number of gaps and options (Tekin, 1996). In this respect, the following mean ranges were specified:

1.00-1.80 Strongly disagree

1.81-2.60 Disagree

2,61-3.40 Neutral

3.41-4.20 Agree

4.21-5.00 Strongly agree

Based on these ranges, those having mean scores between 1.00 and 2.60 were labelled as loaded with "negative attitudes toward English" while those having mean scores ranging from 3,41 to 5,00 were labelled as "positive toward English". On the other hand, those having mean scores between 2,61 and 3,40 were labelled as neither negative nor positive toward English.

To conclude, as a result of the procedures stated above to use the scale, the scale was confirmed with 19 items collected around four factors as a valid and reliable assessor of university prep EFL learners' attitudes toward EFL (RMSEA=0.068; SRMR= 0.012; $\chi2 / df = 1.52$; α = 0.91).

3.3.4. Semi-Structured Interview Form

Taking the commonly cited advantages of the flipped classroom model into account, the researcher, also as the teacher in the class, aimed to conduct a study informing

the concerned bodies about a better EFL practice. In this respect, analyzing the participants' self-reported opinions about their flipped experience and its role in the predictors of EFL achievement such as attitudes and self-efficacy beliefs becomes significant to better understand the issue under the studied setting. With this aim of gaining insights into the participant experiment group students' perceptions about their flipped classroom experience and its impact on their attitudes toward and self-efficacy beliefs in EFL, a semi-structured interview form was developed by the researcher (see Appendix 11). The reason behind the choice of this instrument as a way of data collection was that interview is not only a way to give the participants a chance to reflect on their experiences but also a way to gain knowledge from them (Kajornboon, 2005).

Interviews are held when the researcher asks the participants some open-ended questions so that they "can best voice their experiences unconstrained by any perspectives of the researcher or past research findings" (Creswell, 2012, 218). In this process, also termed as standardized open-ended interview (Patton, 1990), the researcher has a number of concepts, issues and accordingly questions in mind. Consistent with this point, the researcher of the current study reviewed the related literature for the questions asked the participants about their flipped classroom experiences, their attitudes toward and self-efficacy beliefs in EFL. As a result, the researcher came up with a form consisted of ten questions (see appendix 11). Six of these questions sought the participants' perceptions of flipping EFL classroom. While five of them aimed to reflect on main components of the model, one of them relied on asking the learners about the application of the model in other subject areas. Three of the remaining questions aimed to elicit the impact of this model on their attitudes toward and self-efficacy belief in EFL. The remaining question, on the other hand, aimed to provide the interviewee with a chance to elaborate on flipped classroom model in any perspective.

Following this, the draft form was sent to two experts from a language department for the relevance and comprehensibility of its language. The draft form was edited as a response to the suggestions made by the experts. Then after, an expert from assessment and evaluation, two experts from curriculum and education, an expert from educational psychology and another from English language teaching were consulted to comment on the face and content validity of the edited form. Here

similar procedures to those followed in the development of the achievement test were held. In this respect, an adaptation of Lawshe technique was adopted. The reason behind it was not to limit the experts to three classical options as necessary, unnecessary and useful but not necessary. In a form, the experts were asked to circle the number that reflects their opinions best from "1 = completely invalid" to "5= completely valid" to give a statistical value for the content and face validity of the form (see Appendix 12). Based on the number of gaps and options as criteria (Tekin, 1996), it was decided to omit the questions with the mean ranging from 1.00 to 2.60 revising those ranging from 2.61 to 3.40. As a result of employing basic descriptive statistics, none of the items were found to be invalid. In other words, the form was validated with the mean score of 4,8 out of 5 indicating the semi structured form as a completely valid assessor of EFL students' perceptions of flipped classroom model in EFL.

To identify the interviewee, the *multiple variation sampling* was employed. In order to "present multiple perspectives of the individuals", the researcher uses this sampling to include "cases or individuals that differ on some characteristic or trait (Creswell, 2012, 208-209). In this respect, the researcher interviewed with nine experiment group students developing differently from pre- to post- academic achievement test. Therefore, 3 of these students were chosen from those whose scores were the most negatively or least positively changed while 3 other students were chosen from those whose scores moderately changed. The remaining 3 interviewee students were those whose scores were the most positively changed from pre to post-test. Each of the interviews was held in the university in a period of 25-40 minutes. Based on the consent forms obtained from the participant interviewees, the interviews were realized and the researcher audio-taped the whole interviews (see Appendix 13).

As a result of the administration of interview to explore the central phenomenon of the study, the researcher obtained "information rich" data and helped to deepen understanding of the efficiency of flipped classroom model in EFL. (Patton, 1990).

3.4. Data Collection Procedure

This part takes the form of the aim to reflect on the procedures followed by the research to conduct the study based on the permission from Institutional Review of

Board (IBR) with a notice of E.11219 issued on September 9, 2016 (see Appendix 15). Within this respect, in five main areas, in order to elaborate on the aforementioned four folds, specific steps taken and procedures followed during the study including the process of data collection are outlined in the following table:

Table 3.3: Data Collection Procedures

Stage	Step	Time
Conceptual Framework	Review of the related literature	May-June, 2016
Instruments	Developing EFL Achievement Test	1st - 3rd Week, July 2016
	Developing Semi Structured Interview Form	3rd - 4th Week, July 2016
	Organizing EFL Self-Efficacy Belief Scale	4th Week, July 2016
	Organizing Attitudes Toward English Scale	4th Week, July 2016
Pilot Study	Official correspondence with the university for the pilot study	4th Week, July 2016
	Identifying the participants of the pilot study	1st Week, August 2016
	Piloting of the Study	August 2016
	Validity and reliability analysis of the scales	1st Week September 2016
	Reporting the results	2nd Week, September 2016
Implementation of the Study	Pre-testing of data collection instruments	2nd Week, September 2016
	Analyzing the results	3rd Week, September 2016
	Assigning control and experiment groups	3rd Week, September 2016
	Flipped classroom intervention	September-December, 2016
	Post-testing of data collection instruments	1st-3rd Week, January 2016
	Analyzing the results	4th Week, January - 2nd Week, February 2017
	Semi-Structured interviews	2nd- 4th Week, February 2017
	Analyzing qualitative data	March, 2017
Report	Concluding the results and finalizing the study	April - May, 2017
Total	12 Months	May,2016-May,2017

As seen in the table, the data for the study was collected in five main phases in 12 months covering a period from May, 2016 to May 2017. The following parts took the form of attempts to reflect on each step in detail.

3.4.1. Pilot-Testing

The data collected during the study was articulated by 38 students who were qualified to study at English Prep Schools of eight different universities (5 state and 3 private) in the pilot study and 41 students enrolled at foreign language (prep) schools at Gebze Technical University in the main study.

To reach the aim of the current study, the researcher needed to adopt scales that could explore university prep EFL students' attitudes toward and self-efficacy beliefs in EFL. Moreover, the researcher also needed to develop a reliable and valid achievement test that could measure the participant students' EFL performance. To support this quantitative data and enlarge the scope of the study, the researcher also needed to develop a semi-structured interview form to obtain data indicating the learners' perceptions of the flipped classroom experience and its impact on their attitudes toward and self-efficacy beliefs in EFL.

In the current study, as implied above, before the actual study, a great deal of effort was devoted to the administration of pilot study to develop and test the adequacy of data collection instruments. This pilot study was preceded by obtaining the permission to administer the scale in the current study from their developers and informing the participants about the overall aim of the study together with their rights to receive their consent. The pilot of the study started with the application of EFL Achievement Test. Having signed the consent form, the students completed the first part of the test consisted of listening, grammar, reading and vocabulary in a period ranging from 50 to 60 minutes. After a fifteen-minute break, the students were observed to complete writing part in a period ranging from 25 to 30 minutes. Following 20-minute break, the students were requested to complete 34-item EFL Self-Efficacy Belief in English Scale (see Appendix 7). Next, 19-item Attitudes toward English Scale was given to the participant students after another 20-minute break. The data gathered as a result of the administration of these 3 instruments was exposed to elaborate validity and reliability analyses. In other words, the data collected from the administration of the test was exposed to item analysis including

item difficulty, item discrimination and the reliability analyses computed through ITEMAN 4. On the other hand, the data acquired out of the scales formed the basis of the adequacy of the sampling, reliability and validity analyses including CFA and alpha value through SPSS version 21 and through LISREL 8,54 for Windows respectively.

Then, three separate *Independent Samples t-tests* were computed to equate the participants in the control and experiment groups in terms of the pre-test results of *EFL Achievement Test*, *Attitudes toward EFL* and *Self-Efficacy Beliefs in EFL* scales in the pilot study. To anticipate unexpected problems in the treatment of major experimental study, the researcher flipped the classroom for the students in the experimental group by uploading web lectures prepared by the researcher himself on YouTube asking them to watch the video lectures at home before the class to allocate classroom time for interactive group activities. On the other hand, other 19 students in the control group were traditionally following lecture and assignment procedures. Following the flipped classroom treatment that took two weeks, *the EFL Achievement Test*, *EFL Self-Efficacy Belief in English* and *Attitudes toward English Scales* were given to the students as post-tests. This was done not to employ the complete statistics as in the real study but to check time spent between the post-tests and the retention test and anticipate some problems that may come out during the main study.

Following two-week treatment, semi-controlled interview form which was developed by the researcher using an adaptation of Lawshe technique justified and validated with the participation of seven experts from different subject areas in different phases were also held with 5 voluntary students in the flipped EFL classroom from three different achievement groups. This interview consisted of ten questions were held not to produce qualitative data to support quantitative data in the pilot study but to explore any problems that may come out during these semi-controlled protocols and tailor the related procedure in the major study.

3.4.2. Pre-Testing

As implied in the table above, apart from the steps taken to review the related literature, locate and develop the data collection instruments and conduct the pilot study to have advance warning about problems that may come out in the main study,

the data collection procedures were followed in four folds. While the first two quantitative folds included pre and post-test implementation of the aforementioned scales interrupted by the flipped intervention, the last fold included the collection of the qualitative data acquired as a result of semi-controlled interviews with the experiment group students about their flipped classroom experiences. This mixed method structured flow of the data collection procedures of the study cut by the experimental intervention is illustrated in the table below as follows:
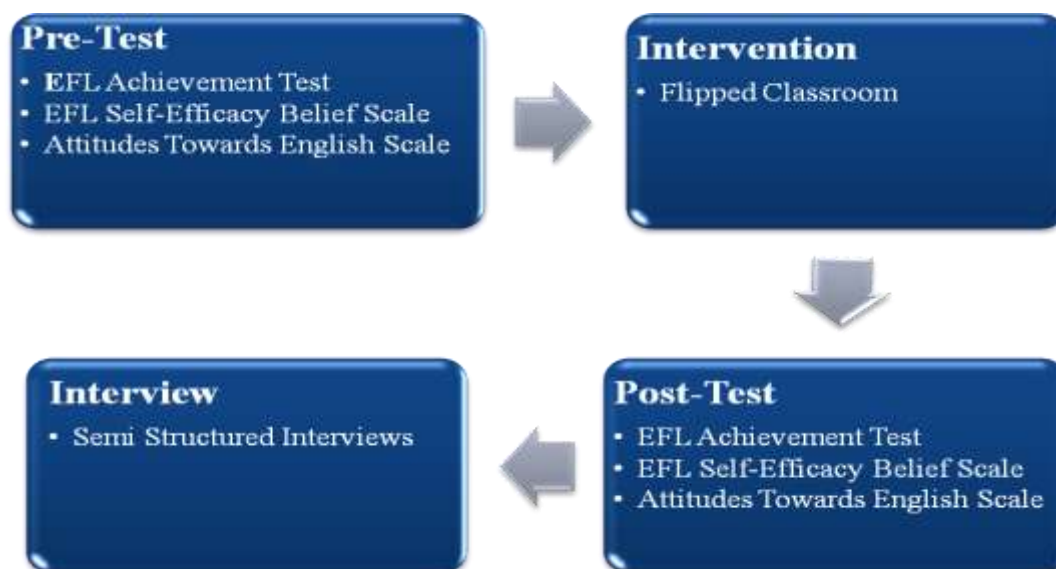
Figure 3.2: Flow of Data Collection Procedures

As seen in Table 3.2, the implementation of the main study started with the pilot testing in September. The pre-testing was initiated by getting the participant students sign the consent forms in the second week of September. Taking the problems that came up with the administration of three data collection instruments consecutively on the same day in the pilot study, the procedures for pre-testing were shaped. In this respect, the researcher asked the participants to take *EFL Achievement Test* and *EFL Self-Efficacy Belief in English Scale* on the first day in the main study. Time ranges determined as a result of pilot study were followed for the application each data collection instrument in this phase. Not to cause any mental fatigue, *Attitudes toward English Scale* was given to the participants on the next day.

Then participants were assigned to control and experimental groups based on the results of three separate *Independent Samples t-tests* computed to equate them in terms of the variables within the concern of the study.

3.4.3. Intervention

Following the procedure where the groups were assigned and equated in terms of three variables of the study, the intervention was initiated. In this respect, to accomplish the internal validity of the design, the courses in both of these groups were led by the researcher himself during the fall term of 2016-2017 academic year. Moreover, the same syllabus prepared and submitted by Gebze Technical University Department of Foreign Languages was followed in both classrooms (see Appendix 16). The instruction lasted 15 weeks from September to December. To realize the aim of the experimental part of the study, the researcher followed Flipped Classroom Model in the experimental class while he adopted traditional lecture-based instruction in the control group. Within this respect, the following two parts were devoted to present the teaching procedures followed in experimental and control groups.

3.4.3.1. Procedures for the Experimental Group

This group of students formed the population of the treatment classroom that would be flipped. Hence, starting with the first week of the term, lecture parts of the lessons in the experimental group were flipped with the relevant homework session outside classroom. The main flow of the procedures in the flipped classroom was shown as follows:
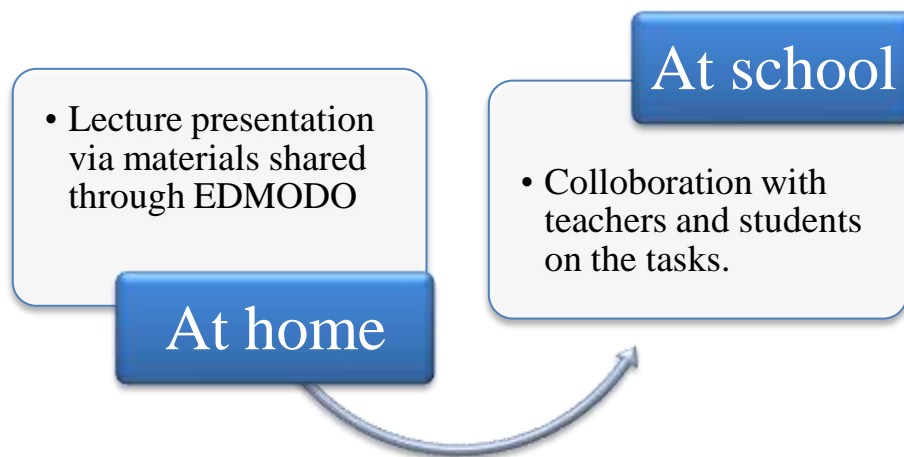
Figure 3.3: Flow of Instructional Procedures for the Experimental Group

At the very beginning of the term, the researcher informed the participants in the experiment group about the flipped classroom procedures both formally through the

consent form and informally in the classroom via PowerPoint presentation after the implementation of the data collection instruments as pre-test. Basically, being a member of this flipped classroom required the students to watch the video lectures created by the researcher or provided by the researcher from different sources such as some web pages where a wide range of materials are offered for online or flipped courses. These were created or selected from already created ones taking the theoretical considerations explained before into account. As a result, the researchers came up with interactive videos with an average run time of 10-15 minutes, *Google docs* or presentations that can flip delivery of the content in the classroom with the assignment at home. These were shared with the students through an online course management system called *Edmodo*. Thanks to this network platform, the participants in the experimental group reached and watched or listened to different sources of materials ranging from videos and audio files to PowerPoint presentations at their houses before their courses at school. Additionally, this network triggered the communication between the teacher and students outside the classroom independent of time and place. The students could easily communicate with one another about a shared material as well as having instant feedback from the teacher.

The researcher also followed the recommendation in the flipped classroom literature that a quiz was a must on the lecture before the classroom session (Zappe et al., 2009). As a result of this, the researcher aimed to get the students adopt the habit of coming to classroom having finished the related short exams prepared by the researcher himself in different formats to ensure that they studied the shared materials. This also made it possible to engage the students with more exercises and more problem solving activities in different skills and spend more time in the classroom on the acquisition of higher order skills. In other words, to accomplish flipped intervention, during the research process, the researcher shared 46 assignments in forms ranging from videos, voiced PowerPoint, *Google docs*, audio files to PDF and word files through *Edmodo* to help the participant student come ready to the classroom. The researcher also uploaded 32 quizzes with different types of items ranging from multiple-choice, fill-in-the blank, true-false to short answer and matching. In addition to track the students who came to the classroom having reviewed the delivered content or not, thanks to the instant online feedback provided by the teacher about the quizzes in different forms, the students completed the

missing parts in the acquisition of the content. The reason behind this was that the quizzes were graded right after they were submitted by the students allowing them to go over their mistakes for partial credits.

The style of the material differed based on the aims of each specific lessons, requirements of each content to be delivered, time limits and researcher's preferences. Yet, all these materials in different forms that address the delivery of each content specified in the syllabus formed the assignment phase of the flipped classroom procedures that is realized independently of time and place outside the classroom.

The virtual flipped classroom that was created on *Edmodo* by registering 22 students (one being the researcher himself to check the quality of the delivered materials as a student) is shown in the screenshot as follow:
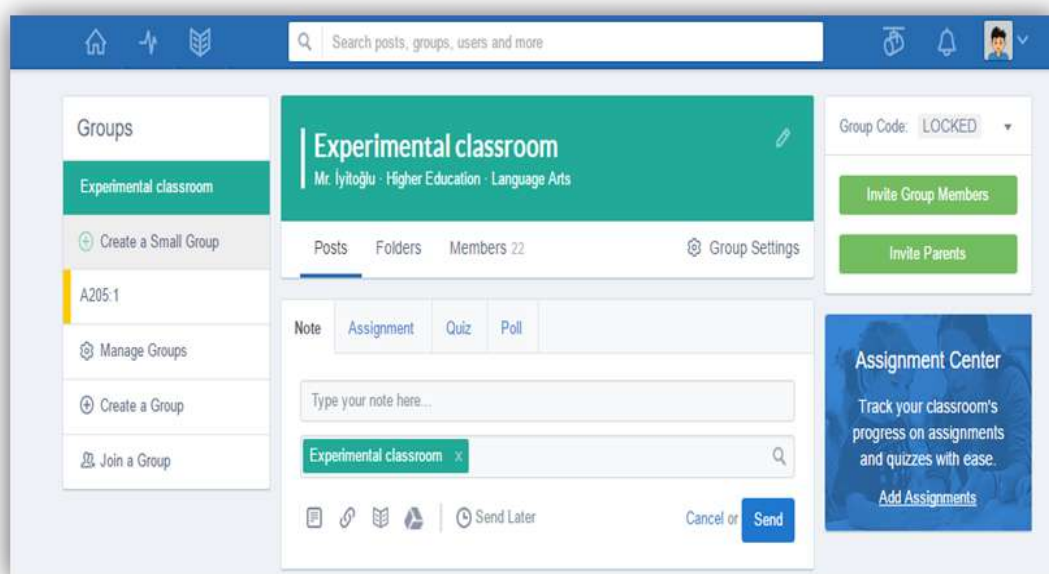
Figure 3.4: Flipped Classroom on Edmodo

In addition to the screenshot of the Edmodo account created for the students in experimental group, some example screenshots of materials shared with the students as assignments (46) and quizzes (32) to check both their understanding and whether they really studied them are also presented below in the following figures:

Figure 3.5: Video Assignment on Simple Present Tense

Figure 3.6: Quiz on Simple Present Tense

In addition to sample assignments and quizzes sent to the participants in the flipped classroom intervention, example screenshots of posted versions of them are also shown in the following figures:

Figure 3.7: Post on If Clauses Assignment
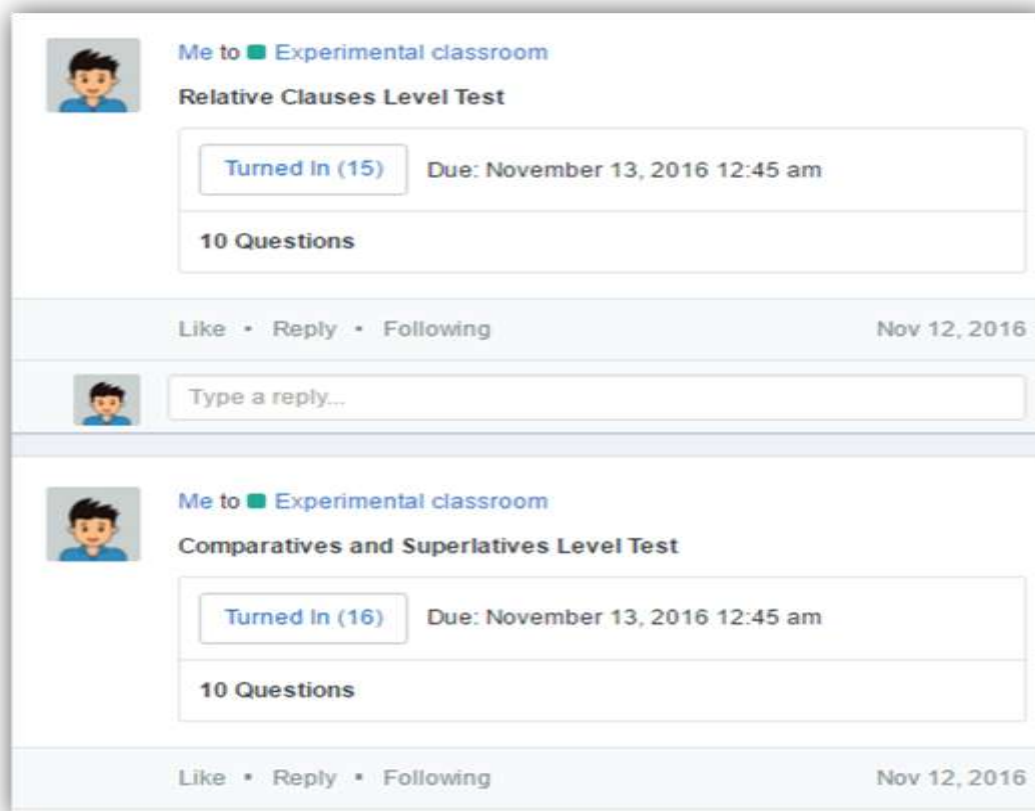
Figure 3.8: Posts on Relative Clauses and Comparative & Superlative Quizzes

After reviewing delivered content as homework, the participant students were exposed to the second part of the flipped classroom sessions in the classroom. This phase served for enhancing higher order skills among students as they were engaged in different interactive group work activities by giving them necessary amount of individualized support in a way to help them master the content. Therefore, the teacher aimed to adopt the role of professional teachers in the flipped classrooms who "continually observe their students, provide them with feedback relevant in the moment, and continuously assess their work by being reflective in their practice, connecting them with each other to improve their trade, accept constructive criticism, and tolerate controlled classroom chaos" (Hamdan, et al., 2013, 6)

Within this respect, the classroom sessions always started with a warm-up activity followed by a whole class discussion about the delivered content. However, the teacher did never give deep lectures about the content. Instead, the teacher briefly reviewed the content through a question and answer session. When necessary, the teacher initiated another session with exercises from the coursebook to reinforce learning. This review session organized in warm-up, whole class discussion or

question and answer activities lasted average time of 10-15 minutes and helped students clarify the question marks regarding what was not understood regarding the content. This generally followed group work activities based on a meaningful problem to help students better reinforce and expand their mastery of the content by having more time to work with their peers through the assistance of the teacher in the classroom. Apart from teaching grammatical structures, this session in the classroom took the form of engaging students in reading, writing, listening or speaking activities based on the content delivered to them through Edmodo. Sometimes, the students were also engaged in fun activities such as Kahoot games. Then, the second and last phase of the flipped classroom finished assigning students what is known as "homework" to cover in collaboration with their peers and the teacher in the classroom. It is important to state that the same homework was assigned to the students in both control and experimental classrooms. The only difference was that the students in the control group had to complete their homework outside the classroom, while for those in the flipped classroom their peers or the teacher were available for help in the classroom During this time, the teacher walked around the classroom and assisted the students individually when needed.

This two-phased flow of procedures held for flipped classroom in the present study was concretely embodied in the following table:

Table 3.4: Flow of Procedures for the Experimental Classroom

Week	Subject	1ST Phase Outside the Classroom	2ND Phase In the Classroom
1	ENGLISH FILE ELEMENTARY UNITS: 1ABC & 2ABC	• PowerPoint Presentation • Video • Voiced PowerPoint Presentation • Teacher-Created Quiz 1 • Teacher-Created Quiz 2	• Warm-up • Whole Class Discussion • Pair Work • Group Work • Dictation • Show and Tell • Sentence Auction • Assignment

Table 3.4 – continue

Week	Subject	1ST Phase Outside the Classroom	2ND Phase In the Classroom
2	ENGLISH FILE ELEMENTARY UNITS: 3ABC & 4ABC	• Voiced PowerPoint Presentation • PDF • Video • Teacher-Created Quiz 3 • Teacher-Created Quiz 4	• Warm-up • Question and Answer • Whole Class Discussion • Problem-Solving Activity • Assignment
3	ENGLISH FILE ELEMENTARY UNITS 5ABC & 6ABC	• PowerPoint Presentation • Google Doc • Videos (2) • Teacher-Created Quiz 5 • Teacher-Created Quiz 6	• Warm-up • Question and Answer • Whole Class Discussion • Story Telling • Assignment
4	ENGLISH FILE ELEMENTARY UNITS: 7ABC & 8ABC & 9A	• PowerPoint Presentation • Voiced PowerPoint Presentation • Google Doc • Teacher-Created Quiz 7 • Teacher-Created Quiz 8	• Warm-up • Whole Class Discussion • Pair Work • Reading For Gist • Teaching Listening Through Songs • Assignment
5	ENGLISH FILE ELEMENTARY UNITS: 9BC & 10ABC & 11AB	• Voiced PowerPoint Presentation • Video • Word File • Teacher-Created Quiz 9 • Teacher-Created Quiz 10	• Warm-up • Whole Class Discussion • Group Work • Dictation • Reading For Specific Information • Bingo Game • Assignment
6	ENGLISH FILE ELEMENTARY UNITS: 11C & 12ABC	• PowerPoint Presentation • Google Doc • Video • Teacher-Created Quiz 11 • Teacher-Created Quiz 12	• Warm-up • Whole Class Discussion • Story Telling • Small Group Discussion • Assignment

Table 3.4 – continue

Week	Subject	1ST Phase Outside the Classroom	2ND Phase In the Classroom
7	ENGLISH FILE PRE-INTERMEDIATE UNITS: 1ABC & 2ABC	• PowerPoint Presentation • Teacher-Created Video • Word File • Teacher-Created Quiz 13 • Teacher-Created Quiz 14	• Warm-up • Question and Answer • Problem-Solving Activity • Bingo Game • Assignment
8	ENGLISH FILE PRE-INTERMEDIATE UNITS: 3ABC & 4ABC	• Voiced PowerPoint Presentation • Audio File • PDF • Teacher-Created Quiz 15 • Teacher-Created Quiz 16 • Teacher-Created Quiz 17	• Warm-up • Question and Answer • Small Group Discussion • Pair Work • Group Work • Scanning • Kahoot Games • Assignment
9	ENGLISH FILE PRE-INTERMEDIATE UNITS: 5ABC & 6ABC	• Voiced PowerPoint Presentation • PDF • Video • Teacher-Created Quiz 18 • Teacher-Created Quiz 19 • Teacher-Created Quiz 20	• Warm-up • Question and Answer • Whole Class Discussion • Brainstorming • Kahoot Games • Group Work • Assignment
10	ENGLISH FILE PRE-INTERMEDIATE UNITS: 7ABC & 8ABC & 9A	• PowerPoint Presentation • Video • Voiced PowerPoint Presentation • Teacher-Created Quiz 21 • Teacher-Created Quiz 22	• Warm-up • Question and Answer • Group Work • Skimming • Taboo • Kahoot Games • Assignment
11	ENGLISH FILE PRE-INTERMEDIATE UNITS: 9BC & 10ABC	• Voiced PowerPoint Presentation • Video • PDF • Teacher-Created Quiz 23 • Teacher-Created Quiz 24	• Warm-up • Question and Answer • Problem-Solving Activity • Assignment

Table 3.4 – continue

Week	Subject	1ST Phase Outside the Classroom	2ND Phase In the Classroom
12	FOCUS 2: UNITS 1 LISTENING Q SKILLS 2: UNIT 1 GREAT WRITING 2: UNIT 1	• Audio File • Google Doc • Video • Teacher-Created Quiz 25 • Teacher-Created Quiz 26	• Warm-up • Whole Class Discussion • Sentence Jumble • Listening for General Meaning • Controlled Writing • Assignment
13	FOCUS 2: UNITS 2 LISTENING Q SKILLS 2: UNIT 2 GREAT WRITING 2: UNIT 2	• Audio File • Google Doc • Video • Teacher-created Video • Teacher-Created Quiz 27 • Teacher-Created Quiz 28	• Warm-up • Whole Class Discussion • Group Work • Listening for Specific Information • Assignment
14	FOCUS 2: UNITS 3 LISTENING Q SKILLS 2: UNIT 3 GREAT WRITING 2: UNIT 3	• Audio File • Google Doc • Video • Teacher-Created Quiz 29 • Teacher-Created Quiz 30	• Warm-up • Whole Class Discussion • Brainstorming • Listening for Specific Information • Assignment
15	FOCUS 2: UNITS 4 LISTENING Q SKILLS 2: UNIT 4 GREAT WRITING 2: UNIT 4	• Teacher-created Video • PowerPoint Presentation • Teacher-Created Quiz 31 • Teacher-Created Quiz 32	• Warm-up • Whole Class Discussion • Listening for the Gist • Semi-Controlled Writing • Assignment

As seen in the table, materials used to deliver the content out of the school and the procedures followed in the classroom to reinforce the mastery of the content in the

classroom were reflected in the detail standing for each week during the fall term of 2016-2017 academic year.

3.4.3.2. Procedures for the Control Group

On the other hand of the intervention, for 15 weeks during the fall term in 2016-2017 academic year, the researcher taught students in the control group traditionally based on direct instructional lecturing. The term "traditional" stood for any teaching style where students were made to learn the content right in the classroom thanks to the teacher presenting it, and practice and reinforce what they were taught in school at home. Within this respect, basically, the students were asked to participate in the lectures carried out in the classroom and do their homework at home.

To serve for the internal validity of the study, the teacher made use of the delivered materials to the students in the flipped classroom via Edmodo while presenting the content in the control classroom. Moreover, the students in the control group were taught following the same coursebook and the syllabus prepared Department of Foreign Languages at Gebze Technical University (see Appendix 16). This was done consciously so that the probable differences between groups could be attributed to differences in instructional methods rather than the materials and factors related to the instructor.

After warming up students at the very beginning of the course, the researcher presented the content through PowerPoint presentations and other videos especially designed for the flipped classroom. This session took almost two thirds of the whole classroom time. Then, to see and refine if the students understood the content or not, the teacher initiated whole class discussions and question-answer sessions. Students were also made to complete some exercises from the coursebook or workbook in the classroom to reinforce their knowledge. Following this, the students were engaged in small group work activities to expand their knowledge about and efficiency of the content. At the end of the courses, the students were assigned homework to be done at home. It is valuable to state that the same homework was assigned to those in the flipped classroom to be done in the classroom.

3.4.4. Post-Testing

After completing the intervention section that lasted fifteen weeks from the 2nd week of September to the end of December, the third phase of the data collection procedures for the actual study started. The third fold of data collection procedures stands for the administration of *EFL Achievement Test*, *EFL Self-Efficacy Belief in English Scale* and *Attitudes toward English Scale* as post-test in the first weeks of January, 2016. In the administration of these instruments, the same procedures as those in pre-test section were followed giving the students same time limits to complete them. Therefore, the first two instruments were applied on the first day after the intervention while the last scale was administered on the next day not to risk the validity of the data.

3.4.5. Procedures for the Interviews

In the last fold of data collection, the researcher aimed to better enlighten the concerned bodies about a better EFL practice by dwelling on the strengths of both types of data to promise information rich data. In this respect, using the ten-item semi-structured form developed and validated by the researcher with the participation of seven experts, the researcher held semi-structured interviews with nine experiment group students developing differently from pre- to post- academic achievement test. In other words, the students were selected from three achievement groups as low, medium and high. Based on the results obtained from pilot study, in a period of 25-40 minutes, the interviews were held at Gebze Technical University and audio-taped with the permission of the interviewee in the last two weeks of February, 2016.

As a result of the collecting all the necessary quantitative and qualitative data in four folds, the actual study was finalized by carrying out the relevant analysis and reporting the result from March to May, 2017.

3.5. Data Analysis

All the quantitative data was analyzed in six folds using Statistical Packages for Social Sciences (SPSS) 21 for Windows, LISREL 6 for Windows and Item and Test Analysis Program –ITEMAN *4*. While the first two folds stand for the analyses of the data obtained in the pilot study to develop, test and confirm the data collection

instruments, the last four folds cover the analyses of the data in the main study to answer the research questions. This part briefly indicates data analysis methods followed in each fold.

In the first fall of data analysis, the data acquired as a result of the administration of *EFL Achievement Test* in the pilot study was exposed to item level and reliability analyses. In this respect, *ITEMAN* was used to compute item analyses including item discrimination (Rpbis) and difficulty (P) indexes and those of the test as a whole through mean, median, standard deviation. *ITEMAN 4* was also used to compute the reliability coefficients through KR_{20} to stand for the index of internal consistency of the test as a whole and its parts since they include dichotomous items. In this fold of the analysis, the item and reliability analyses were covered separately for both sections as listening, grammar, reading and vocabulary and the whole test. However, the stated analyses were not carried out for writing part due to its open-ended nature. To accomplish the reliability of scoring for writing part, three EFL teachers were asked to follow the guidelines in the analytic rubric provided in the detailed answer key. In the rubric, four areas (grammar, vocabulary, content and organization) were shown as the basis of assessing students' EFL writing performance analytically and separately out of 5 for each of four areas. Participants' EFL writing performance was measured through the mean of three independent EFL assessors. In this respect, reliability for assessing writing performance was realized through the mean scores of three assessors based on analytic rubric while validity was ensured consulting 5 EFL experts' views on the table of specifications for writing.

The second fold of the analysis was devoted to testing and confirming the statistical relevance of *EFL Self-Efficacy Belief Scale* and *Attitudes toward English Scale* to the sample of the study. These required analyses were carried out in four steps. The first step was to compute KMO *& Bartlett's Test of Sphericity* to test the adequacy of the sampling to measure intended analyses in the pilot study. Following this, the researcher computed the *Kolmogorov-Smirnov Test* to assess the normality of the data to choose the best estimation method followed in CFA. In the third step, the researcher used LISREL 6 for Windows to perform confirmatory factor analyses (CFA) in order to explore if the scales are valid measures of the related constructs. The researcher took the last step to run reliability coefficient analysis using *SPSS 21*

for Windows with a significance level of 0.05 to elicit Cronbach's alpha (α) value since it describes internal consistency of a scale with likert items.

In the third fold of the analysis procedures, the researcher aimed to form the basis on which he could adopt the best data analysis methods. In other words, this fold covered the attempts made to determine the type of the data obtained as a result of the administration of pre- and post-tests since there are two types of data, parametric and non-parametric each requiring different types of analysis. Therefore, the researcher checked data in terms of the number of participants, measures of central tendency, normal distribution curves and further ran *Kolmogorov Smirnov tests* to test the normality of the distribution. In this way, the researcher aimed to compute relevant parametric or non-parametric tests on the data and ensure the validity of the conclusions drawn about the sample based on valid findings.

In the fourth step of the analysis, the results of analysis that indicate the data distributed normally were taken as the basis. On this ground, parametric tests were employed in order to answer the research questions of the study. Correspondingly, the pre-test scores of *EFL Achievement Test, EFL Self-Efficacy Belief Scale* and *Attitudes toward English Scale* were analyzed using *Independent Sample T Test*. Based on the results of this test conducted to equate the characteristics of the participant EFL students, two groups of students among whom there was no significant difference were assigned as control and experiment groups.

The following fifth fold of the analysis was devoted to explore the impact of flipped classroom model on EFL learners' scores of *EFL Achievement Test*, *EFL Self-Efficacy Belief in English* and *Attitudes toward English Scales*. To compute adequate tests to produce sound results, the researcher reviewed the related literature to decide on the data analysis method. As a result, it was found out that five main methods were followed and suggested to explore the impact of experimental intervention on dependent variables (Balcı, 2005; Büyüköztürk, 2007a and 2007b; Creswell, 2012; Karasar, 2007):

1. For each group (control and experimental), by computing the percentage increases in pre-test and post-test scores, the mean increases are compared.

2. Controlling the effect of pre-test scores as covariates, Analysis of covariance (ANCOVA) is run to examine the differences in post-test scores.

3. *Multiple Regression Analysis* is run to predict the value of dependent variable based on the known values of independent variables such as pre-test scores and groups.

4. First, the pre-test scores are compared, and if there is no significant difference between them, only the post-test scores are used to test the differences between the means.

5. Depending on the number of grouping variables and measures, appropriate form of Analysis of Variance (ANOVA) tests ranging from One-way, Two-way, MANOVA to Repeated Measures ANOVA is suggested for use in data analysis in experimental designs.

In this respect, depending on the variables and research questions of the present study that probe the difference in the mean of pre-test and post-test scores within and between groups, *Independent Samples T* and *Paired Sample T Tests* were adopted as the main data analysis method.

Although it was partly suitable to be used with the collected data, ANCOVA was not employed for a number of reasons. First of all, this test was found to be used in only studies that seek the impact of intervention on post-test scores controlling the effect of pre-test scores but not the difference within groups (Büyüköztürk, 2007b). In this respect, ANCOVA is basically used not to compare means but to both compare regression lines and indirectly assess the effect of covariates (Myoung, 2009). However, the present study aimed to shed lights on the differences between mean pre-test and pos-test scores within control and experimental groups.

Secondly, ANCOVA in experimental studies is based on null hypothesis (H_0) that there is not a significant difference between pre-test scores between the groups. This happens when the researchers do not reach the related data to statistically accept the the null hypothesis which results in insignificant impact of a variable assigned as covariate on a dependent variable or lose data during the intervention causing them to reject H_0 (Büyüköztürk, 2007b, Myoung, 2009). As a result, basically the pre-tests are assumed to be equal as a result of which they are assigned as covariates and the differences in post-test scores are assessed (Balcı, 2005). However, it was not needed to control the pre-test scores as covariates in the current study since, thanks to *independent sample t tests* explained in the former fold of data analyses, they were

already statistically equated at the very beginning of the study before the intervention. Hence, there seemed no reason to justify the control of pre-test scores in order to explore the significant impact of intervention on post-test scores. On the other hand, despite the differences in their mathematical correctness and power to reduce error, t tests and ANCOVA produce similar results when there is no significant difference in the pre-test scores between the groups (Büyüköztrük, 2007b).

In addition to them, the study did not also seek the joint impact of or interaction between independent variables on dependent variables as in the case of ANOVA. As a result of no interest in the interaction, the present case can still be regarded as a pairwise comparison, which eliminates any inflated type-I error caused by multiple comparisons (Myoung, 2009).

In this study, in order to reach the results that enlighten the readers about the impact of flipped classroom treatment on three independent variables of the study, the researcher did not run factorial analysis, also known as MANOVA and Two-Way MANOVA, either. These tests are used to explore the effect of two independent variables on more than one or more dependent variables simultaneously protecting the researcher from coming up with type 1 error (Büyüköztürk, 2007). However, the fact that the researcher equated the groups in terms of the variables indicated by the results of pre-tests as well as lack of interest in the joint impact of independent variables or the interaction between them on the dependent variable triggered the researcher to employ t-tests to gain insights into the impact of flipped model on EFL achievement, attitudes toward and self-efficacy beliefs in EFL independent of one another.

Dwelling on the reasons and review of literature stated above, *Independent Samples T* and *Paired Sample T Tests* were run to explore whether the experiment and control group EFL learners' mean scores of EFL Achievement Test as a whole, its sub-sections (grammar, listening, reading, writing and vocabulary) and its administration as a retention test differ significantly between or within groups. The same procedures were repeated for the significant difference in the mean scores of *Self-Efficacy Belief in English* and *Attitudes toward English Scales* between and within groups. In more specific terms, separate *Paired Sample T Tests*, also known as *Dependent Sample T Tests*, were run to indicate whether pre- and post-test mean scores of experiment and

control groups related to each of the independent variables (academic achievement, attitude and self-efficacy belief) differ significantly within themselves. This test was chosen since it compares two mean scores of the same group in order to indicate any kind of significant difference (Balcı, 2005). Then after, *Independent Sample T Tests* were administered in order to evaluate the differences between control and experiment group EFL learners' mean scores of *Achievement Test, Self-Efficacy Belief in English* and *Attitudes toward English Scales* conducted as post-tests. As a result of the administration of these two tests, the impact of flipped classroom model on the difference within and between group variations were deeply analyzed.

The last fold of the analysis procedure took the form of the attempt to enrich and support the results of the study by gaining insights into the impact of flipped model on EFL achievement, attitudes toward and self-efficacy beliefs in EFL. This data was analyzed using content analysis method (Creswell, 2012; Patton, 1990). In this respect, the qualitative data gathered as a result of interviews was analyzed manually following content analysis procedures including the steps to transcribe the data, code the data according to the pre-determined themes by the researcher to stand for each of the questions in the interview form.

After transcribing the recorded interviews, the process of forming the codes out of the participants' responses started. For the accuracy of these codes, the interviewees were asked to reflect on them. These codes were classified based on 9 themes pre-determined by the researcher to stand for each of the questions in the interview. However, the last question was an open-ended one to allow the interviewee to elaborate on their flipped experience in any way. Therefore, no theme was determined in advance to stand for it. This process continued till the saturation was achieved. Once all the codes and themes were verified, the frequencies and percentages were calculated using SPSS 21 for Windows. The interpretations and comments on EFL students' reflections on the role of flipped classroom in their EFL achievement, attitudes toward and self-efficacy beliefs in EFL were grounded in these themes and codes.

Reliability and validity are conceptualized in different terms such as *trustworthiness, rigor, quality* in qualitative studies (Golafshani, 2003). These are ensured in different ways in qualitative research paradigms. First of all, the researcher is defined by Patton (1990) as the instrument in the qualitative research. Therefore, the researcher

of the study aimed to ensure the reliability of the results by thickly describing the setting, every step of qualitative data collection and analysis procedures. On the other hand, to further validate the accuracy of the results, the researcher employed triangulation and member checking (Creswell, 2012; Golafshani, 2003). Through triangulation the researcher drew on the information obtained from both surveys and the interviews. Moreover, the researcher also employed member checking and asked the interviewee to comment on the accuracy of the codes and the fairness of the interpretations. While some of the interviewees were sent e-mails, some of them were informed in face to face meetings. As a result, the researcher obtained accurate and trustable results regarding the participants' experience of flipped EFL classroom and its impact on their attitudes and self-efficacy beliefs.

3.6. Validity Threats

The primary aim of the study is to delve into the cause and effect relationship between the dependent and independent variables of the study through a mixed method design. In this respect, the study aims to generalize the results, which are only attributed to the independent variables not other factors, to the population at large. This is only possible when the whole research process is clearly set and the results are acceptable and consistent to other researchers (Yıldırım, Şimşek, 2006). As a result, this aim drove the researcher to deal with the threats to internal and external validity of the present study by following the most recommended methods in the literature (Guba, Lincoln, 1982; Yıldırım, Şimşek, 2006):

- To ensure the credibility of a study, the researchers are recommended to clearly state the whole research process with well-constructed research questions (Yıldırım, Şimşek, 2006). In this respect, the researcher designed the study based on the research questions evaluated, edited and confirmed by experts.

- To serve for the internal validity, the researcher needs to show with confidence that the changes in dependent variables are caused by the changes in the independent variables (Creswell, 2012). Therefore, the researcher of the study aimed to clarify the study was free from the influence of confounding variables by leading the courses in both groups following the same curriculum. On the other hand, to attribute the changes in the dependent

variables to the changes in the independent variables, in addition to equating two groups at the very beginning of the study, the researcher made use of the delivered content in the control group in the classroom.

- External validity refers to the extent to which the results of a study can be generalized (Yıldırım, Şimşek, 2006). Therefore, the researcher included participants who are representative of the population at large by randomly assigning intact groups as control and experimental groups. On the other hand, the interviewees were specified by employing the *multiple variation sampling* to "present multiple perspectives of the individuals… that differ on some characteristic or trait (Creswell, 2012, 208-209).

- To further validate the accuracy of the results and interpretations, the researcher consulted experts' views while developing, testing data collection instruments and analyzing the results.

- Moreover, the researcher received feedback from the interviewee to confirm the fairness of the interpretation regarding qualitative results. In addition to that, the researcher applied peer debriefing to work with impartial colleagues or experts analyzing the data. In this respect, the consistency of the results is ensured through peer debriefing and member checking.

- Consistency of results is also recommended to be achieved through triangulation (Golafshani, 2003). Hence, the researcher enriched the data by drawing on the information obtained from different sources including both surveys and the interviews to judge the quality of the results.

- As set by Patton (1990), the researcher himself is the instrument in the qualitative research. Therefore, the researcher of the study thickly described the setting, every step of qualitative data collection and analysis procedures to attain the credibility of the researcher's interpretation.

- To ensure trustworthiness, in addition to thick description, the researcher exemplified the interviewee's responses directly through quotation and avoided data loss by audio-taping the interviews based on their consent.

- Long-term interaction between the researcher and the participants improves quality of the results (Yıldırım, şimşek, 2006). In this respect, the data for the study was collected for the whole process of 2016-2017 fall term.

- All the data was analyzed objectively by consulting different expert evaluations.

Based on all these methods followed, the researcher aimed to ensure internal and external validity of the study. The related procedures were deeply elaborated in the sub-sections of the study ranging from *research questions, design, participants* to *data collection, analysis and results.* In this respect, the results that were achieved based on the findings of the study, which were collected from multiple sources and cross checked thanks to the feedback from the experts, colleagues and the participants, are supposed to be only because of the changes in the independent variables of the study and are qualified enough to generalize to the population at large. In other words, within its limitations, the present study is assumed to enable a valid comparison between groups.

3.7. Pilot Study

Pilot study stands for a small scale study designed to evaluate such facets of the real study as feasibility, cost, time, reliability and validity of the data collection instruments (Van Teijlingen, Hundley, 2002). In this respect, in order to overcome the potential problems that may come out during the implementation of the actual study and, accordingly, to improve its quality and efficiency, researchers conduct pilot studies with a small group of participants who are representative of the target samples.

Within this perspective, the researcher conducted the pilot of the current study with 38 students who were qualified to study at English Prep Schools of eight different universities (5 state and 3 private) in Istanbul in 2016-2017 academic year (male= 24, female= 14). In an effort to compose the sampling of the pilot study representative of the sampling of the main experiment, the researcher employed convenient sampling method. This type of sampling was administered by the researcher who invited his former students from high school taking university entrance exam in 2016 since they were "willing and available to be studied"

(Creswell, 2012, 145). These participant students were selected since they were representative of the sampling of the major study enrolled at Department of Foreign Languages, Gebze Technical University sharing some characteristics to the concern of the study. This relevance was rooted in and backed up by the results of *Undergraduate Placement Examination (LYS)* designed by *Student Selection and Placement Centre (OSYM)* for the admission of the students to higher education in Turkey. Based on Simon's (2011) remark that 10 or 20 % of the sample size of the actual study is reasonable, a participation of 38 prospective university prep EFL students can be justified.

The participant students of the pilot study were asked to voluntarily take part in the pilot study at Gebze Technical University in August and sign the consent form where they were informed about not only the aim of the study but also their rights (see Appendix 13). Justifying the participation on voluntary basis, this pilot study was fulfilled, excluding the retention test, in the period of two weeks in August 2016.

Pilot studies are reported to address a number of different issues (Van Teijlingen, Hundley, 2002; Simon, 2011). Possessing characteristics as a limited and trial version of the planned study, it serves to reveal any unexpected problems during the research study including procedures run to test the adequacy of the data collection instruments.

Within the perspective of conducting pilot studies stated above, the researcher of the present study aimed to test the adequacy of the data collection instruments as specified in the previous parts. On the other hand, the research also aimed to "identify potential practical problems in following the research procedure" (Van Teijlingen, Hundley, 2002, 33). More specifically, the primary aim of this pilot study was to administer item analysis on the *EFL Achievement Test* including item discrimination, item difficulty. On the other hand, the pilot test served to test *EFL Self-Efficacy Belief in English and Attitudes toward English Scales* for the reliability through computing cronbach's alpha value, KR_{20} and validity through employing CFA (Confirmatory Factor Analysis). The reliability and validity values of the scales used in the study were previously reported to be within acceptable ranges by the justification of cronbach alpha values and fit indices. The developers of the scales, which were used to collect data in the present study, computed both EFA (Exploratory Factor Analysis) and CFA to explore and confirm the underlying factor

structure of the stated constructs as "attitudes toward EFL" and "self-efficacy beliefs in EFL". In this respect, to produce sound outcomes, the researcher administered only confirmatory factor analyses (CFA) and reliability coefficient analyses to test and confirm the adopted scales as a valid and reliable measure of the stated constructs within the sampling of the pilot study representative of those in the major study.

The pilot of the study started with the application of *EFL Achievement Test* developed by the researcher based on the justified critical objectives and functions that the students were to acquire during the term at the Department of Foreign Languages at Gebze University. Following signing the consent form documenting their participation on the voluntary basis, the students completed the first part of the test consisted of listening, grammar, reading and vocabulary in a period ranging from 50 to 60 minutes. After a fifteen-minute break, the students were observed to complete writing part in a period ranging from 25 to 30 minutes. By providing feedback on research procedures, the pilot study indicated the convenience of the allocated time, adequacy of wording and form, comprehensibility of the instructions and place regarding the EFL Achievement Test and its administration. The data collected as a result of the administration of this test in the pilot study was exposed to item analysis including item difficulty, item discrimination and the reliability analyses computed through *ITEMAN 4*.

Following the administration of *EFL Achievement Test*, the students were asked to have 20-minute break. Then, the students were requested to complete 34-item *EFL Self-Efficacy Belief in English Scale* (see Appendix 7). This scale was originally developed to explore the high school EFL learners' self-efficacy beliefs in English. Therefore, the pilot study becomes more significant to test the relevance of the scale to the participants of the real study. As a result of the implementation of the scale, it was seen that students spent 10 to 20 minutes to complete it. In this respect, in the major experimental study, it was decided to allocate maximum 20 minute for the completion of *EFL Self-Efficacy Belief in English Scale.* Moreover, the pilot study justified the break time between two data collection instruments, wording and the form of the scale. Furthermore, the data obtained as a result of this procedure, formed the basis of the adequacy of the sampling, reliability and validity analyses including

CFA and alpha value through *SPSS version 21* and through *LISREL 8,54 for Windows* respectively.

Next, 19-item *Attitudes toward English Scale* was given to the participant students after another 20-minute break (see Appendix 9). Since the validity and reliability analyses of this scale were already handled with university prep EFL students, this procedure of the pilot study was expected to empower the relevance of the scale to the sampling of the study. It was explored as a result of this step of the pilot study that it took 15 to 20 minutes for the students to complete the scale. The pilot study enlightened the researcher about the problem that came up with the administration of three instruments consecutively. It was observed that some students complained about the mental fatigue as a result of following such a procedure. In this respect, the researcher did not want to put the validity and the reliability of the data into risk and, therefore, decided to employ *Attitudes toward English Scale* on the second day of data collection for the actual study. The data gathered at the end of this procedure was also analyzed for the adequacy of the sampling, the reliability and validity through CFA employing *SPSS version 21* and *LISREL 8,54 for Windows respectively.*

As a result of administering three data collection instruments, the researcher set to compute simple descriptive analyses including mean scores, percentages, standard deviations regarding scores of the aforementioned test and scales. As a result of this procedure, the researcher aimed to come up with two groups of students equated in terms of the variables within the interest of the study. The reason behind this is necessity of equating the groups in quasi-experimental studies where random assignment is not possible since such a difference between the groups in terms of measured variables pose threats to internal validity of study depriving it of the ability to produce valid results (Büyüköztürk, 2007; Creswell, 2012). In this respect, three separate *Independent Samples t-tests* were computed to explore if there was any significant difference between two groups of students (N= 19; N= 19) formed in terms of the pre-test results of *EFL Achievement Test*, *Attitudes toward EFL* and *Self-Efficacy Beliefs in EFL* scales in the pilot study. The results of the tests were shown in the following table.

Table 3.5: Independent Samples T-Tests for Equating the Groups in the Pilot Study

Factor	Group	N	$\bar{x}$	S	$Sh_{\bar{x}}$	t	Df	p	D
EFL Achievement	Control	19	47,10	12,431	2,852	0,480	36	,634	----
	Exper.	19	44,84	16,374	3,756				
Attitudes	Control	19	3,1676	,35260	,0808	-0,772	36	,445	----
	Exper.	19	3,2410	,21794	,0500				
Self-Efficacy	Control	19	2,7510	,56222	,1289	1,226	36	,228	-----
	Exper.	19	2,5250	,57396	,1316				

As shown in Table 3.5, *Independent Samples T-Tests* were employed to find out if created two groups of EFL students differ in terms of the concerned variables of the study. The results indicated there was no significant difference between two groups of students regarding their scores of three variables (t = ,480; -,772; 1,226; p > .05). In other words, it was explored that two groups were assumed to be similar based on their pretest results.

As a result of equating procedures, the distribution of the participants and their genders according to the control and experiment groups in the pilot study were shown in the following table.

Table 3.6: Descriptive Statistics for the Experimental & Control Groups in the Pilot Study

Group	Gender	N	Total
Experiment	Male	11	19
	Female	8	
Control	Male	12	19
	Female	7	

As seen in the table, 38 students participated in the pilot study. While 19 of them were in the experimental group taught following flipped learning model, the other 19

students in the control group were taught traditionally following instruction and assignment procedures.

The pilot study serves to enlighten the researchers about the issues related to the process and management of the study (Thabane et. al., 2010). Correspondingly, the researcher piloted the flipped learning treatment of the study to explore whether some problems could emerge that would disrupt the treatment itself for two weeks in August, 2016. Therefore, equation of these two groups based on the results of three instruments was employed to anticipate unexpected problems in the treatment of major experimental study. Therefore, in order to see the practical troubles regarding the implementation of the video content in advance, the researcher made two simple web lectures on simple present tense ready for use in the pilot study. The researcher took the cited fact in the related literature that students preferred shorter and live in-person video lectures more into account (Bishop, Verleger, 2013). As a result, these lectures involved 10 to 15 minute video recordings of the researcher as the teacher presenting the session through power point presentation and another teacher teaching simple present structure. The researcher used these two different types of video content to check students' perceptions about them and their adequacy with them. The screenshots of these web courses were shown in the following figures.

Figure 3.9: Video on Simple Present Tense

CHAPTER 3
Using the Simple Present

CONTENTS

Figure 3.10: Voiced PowerPoint Presentation on Simple Present

The researcher informed the participants in the experiment group about the flipped procedure both in the consent form and in the classroom after the implementation of the data collection instruments. Accordingly, these lectures were uploaded for them on YouTube and the participants in the flipped class were asked to watch the video lectures at home before the class. The class session was allocated for interactive group activities.

The pilot of the flipped treatment indicated that it was hard to track the students who came to the class by watching or not watching the video. When students were asked whether they watched the video or not they were stating that they had done so. However, by probing questions, the teacher explored that some of them had not watched the video. Since watching the pre-class video lectures is one of the main components of flipped classroom model, it was a must to ensure that students watched the videos before they came to the class. Hence, by taking the problems seen in the pilot study into consideration, the researcher understood that it was a must to use a specific program that would track and record the students watching the video lectures. As a result of being informed about this unanticipated problem of the treatment thanks to the pilot study, the researcher took the administrations followed

in similar studies on flipped learning into account (Ekmekçi, 2014; Boyraz, 2014). In this respect, the researcher resolved to use *Edmodo*, a course management system where teachers can collaborate, share content with their students, assign till the due time and check their homework and developments. In this web based platform, it is possible for the teacher to manage or control a classroom by uploading any type of data for the students, who are accepted by the teacher to the group, to access. On the other hand, the students can also communicate with their teachers, see their grades immediately and track their own improvement during the term.

Thanks to the feedback provided by the pilot study, the researcher tailored the research procedures for the major experimental study. In this respect, the teacher, the researcher himself, decided to register the participants in the experiment group to a web based classroom created by *Edmodo*. It was decided to register the students to the classroom by a code sent by the teacher. In this respect, it was intended to flip the lecture and homework sessions of EFL lessons in the experiment group. Using their *Edmodo* account, these students were asked to watch the video lectures created by the researcher himself or provided by the researcher from different sources before coming to school allowing the teacher to spend more time on the acquisition of higher order skills through more open-ended and problem solving activities in the classroom. To ensure that the students watched the online sessions, each session was followed by a brief online quiz prepared by the researcher employing the related application in *Edmodo*. It allows teachers to prepare online quizzes with multiple choice, fill in the blanks, matching, true-false and short answer items. It allows students to go back and forth on the video to answer the questions. However, once the students have submitted the whole test they cannot take the test again. The platform provides immediate feedback to the students on their performance of the quizzes. It also enlightens the teachers about the students' performance on each of the items in the quiz. Therefore, this feature of the application provides teachers with a precious input to start every following classes and tailor the course to best serve for students. This program, in this way, was expected to serve as a trigger and a solution for the students to come to the class having watched the videos. On the contrary to the experiment group, other 18 students in the control group were taught following a traditional approach based on lecturing during the pilot study. On the other hand, it is

important to state that the same syllabus and coursebook was followed by the researcher with two groups of EFL learners during the pilot study.

Following the flipped classroom treatment that took two weeks, taking the problems come out in the pre-test session into account, the participant students were asked to take *the EFL Achievement Test*, *EFL Self-Efficacy Belief in English* and *Attitudes Toward English Scales* as post-tests. During the post-test implementation, students did not seem to have or cause any kind of problem. This procedure was held in the pilot study not to employ the complete statistics as in the real study but to check time spent between the post-tests and the retention test. This was important to decide on the ideal interval between the end of the treatment and the retention test. To learn it, the researcher made a modification in the ultimate test reached on the data from pre-testing and put a box beside each of the questions. Students, taking the test as retention test, were asked to put ticks into the boxes for the questions they remembered. Based on the responses obtained, it was understood that most of the questions (88 %) were not remembered by the participants. This indicated two-week interval is enough to administer the *EFL Achievement Test* to explore the impact of flipped EFL classroom on the retention of EFL achievement.

At the end of two-week treatment, semi-controlled interviews were also held with 5 voluntary students in the flipped EFL classroom from three different achievement group as low, medium and high. This interview consisted of ten questions justified and validated with the participation of seven experts from different subject areas in different phases as specified in part 3.4. These interviews were held individually at university at the consensual time. These interviews were held not to produce qualitative data to be analyzed to support quantitative data in the pilot study. However, based on the function of pilot studies to find out unexpected problems that may hinder the research procedures (Van Teijlingen, Hundley, 2002), these interviews were held to explore any problems that may come out during these semi-controlled protocols and tailor the related procedure in the major study. As a result of this procedure, the interviews were found to last 25 to 40 minutes. During the interview, it was seen that the interview questions were clearly understood by the interviewee and the place where the interviews were realized and the anticipated time allocated for each of them were found to be adequate.

To sum up, the pilot study drew the framework of the current study reflecting the probable problems that may come out in the study and destroy its logical flow. Based on these reflections, the necessary steps were taken by the researcher to ensure both quality and efficiency of the study.

In the following three sections, based on the data generated by the pilot study, the results of item analysis and KR_{20} for *EFL Achievement Test*, cronbach's alpha and CFA analyses employed to calculate the internal consistency and validity of *EFL Self-Efficacy Belief in English* and *Attitudes Toward English Scales* were presented in detail in separate parts standing for each of these three instruments.

3.7.1. Pilot of EFL Achievement Test

EFL Achievement Test was developed by the researcher to elicit the progress of the students at Prep School in Gebze Technical University. Adopting a skill based approach, the test was prepared in such five areas as listening, grammar, reading, vocabulary and writing. Due to its nature that may cause problems in testing and students' limited ability to produce the language, speaking was not included in the EFL Achievement Test. In the pilot study employed with the participation of 38 students qualified to study at English Prep Schools of eight different universities (5 state and 3 private) in Turkey in 2016-2017, the test was administered. The data obtained as a result of this was exposed to item analysis including item difficulty, item discrimination and reliability analysis for computing KR_{20} through *ITEMAN 4*.

In the following parts, the item and reliability analyses regarding listening, grammar, reading and vocabulary sections were covered separately. However, due to its open ended and creative nature, the writing part was excluded and assessed by three EFL teachers following the guidelines in the analytic rubric provided in the detailed answer key. In the rubric, four areas (grammar, vocabulary, content and organization) were taken into consideration to assess students' EFL writing performance analytically and separately out of 5. Participants' EFL writing performance was ultimately produced capturing the mean of three EFL assessors. As a result, reliability was assessed through the mean scores of three assessors based on analytic rubric while validity was ensured consulting 5 EFL experts' views on the harmony between the objective and the item regarding writing.

3.7.1.1. Pilot of Grammar Part

In the grammar section of the test, students' performance on 7 critical objectives was aimed to be evaluated. Due to item analyses that may result in removing some of the items due to their poor qualities, the researcher aimed to initially include at least two items to test each of the objectives in the table of specifications. 21 items in 3 different sub-sections were exposed to item-level statistics. The results were presented in detail as follows:

Table 3.7: Summary Statistics for the Piloted Grammar Section

Score	Items	Mean	S	Min Score	Max Score	Mean P	Mean Rpbis	KR$_{20}$	SEM
Scored Items	21	11,872	4,501	1	21	0,587	0,401	0,74	1,87

As specified previously, while deciding on the items to be included in the ultimate test, the acceptable discrimination and difficulty ranges put forward by Turgut and Baykul (2012) as a composition of different views in the literature were adopted. In this respect, two main concepts attract attention in Table 3.7: Item difficulty and discrimination. However, first of all, as seen in the table, there was a spread of scores, with a s = 4.501 and a range of 1 to 21. Mean of items difficulty (P) in this part stands for the proportion of the test-takers who correctly answered the items. Mean of items difficulty in the grammar section indicates that items are moderately difficult with *P* = 0.587. On the other hand, item discrimination (Rpbis), in the 0 to 1 range, refers to the power of the items to differentiate between participants with high and low scores. As seen in the table above, the items have positive Rpbis at 0,401. As a result of following the framework provided by Turgut and Baykul (2012) to decide on the qualified items, the 21-item grammar section was found to have mean of item discrimination and difficulty within acceptable ranges.

Kuder-Richardson Formula 20 (Kr$_{20}$) stands for the index of internal consistency of the test with dichotomous choices. In other words, it is applicable when each item of the test either true or false. It ranges from 0.0 to 1.0 indexing the consistency of the

measurements about students' level of knowledge of the content assessed by the test. As a special case of Cronbach's α, there is no single accurate answer for acceptable ranges of KR_{20}. Yet, it is generally determined in the related literature that values greater than or equal to .70 indicates acceptable reliability of measurements for teacher-made tests (Cortina, 1993; McGahee, Ball, 2009; Sabri, 2013.)

Within this perspective, as seen in the table, Kr_{20} of 0,74 indicates that this piloted part, as a whole, is a reliable assessor of EFL performance on the specified objectives. Standard error of measurement (SEM) stands for an index of the expected variation in observed scores of the test-takers due to measurement error. In other words, as set by Harvill (1991), SEM is calculated to interpret the difference in the obtained and true scores. On an achievement test, a SEM of $\pm$ 2 indicates that the researchers or teachers can be 95 % confident that true scores fall between 2 x SEM + observed scores and 2 x SEM – observed scores (Harvill, 1991). In this respect, as seen in the table, SEM of 1,87 indicates that this part is able to accurately estimate true scores of the test-takers in the range from +2 x 1,87 to -2 x 1 with 95 % accuracy.

The statistics presented above only provide information for the entire grammar section. In this part there are three sub-sections assessing the participant students' acquisition of 7 different objectives. These 7 objectives are assessed through 21 items. However, the researcher also aimed to use detailed statistics for evaluating individual items for each of these 3 sub-sections. In this respect, for quality control of the first sub-section of the whole grammar section, which was made of 11 multiple choice items, item-level analyses were administered and the results were presented as follows:

Table 3.8: Item Statistics of the Piloted Multiple-Choice Grammar Items

Item	N	P	Total Rpbis	Alpha w/o	Option	N	Prop.	Rpbis	Mean	S
1	38	0,639	0,630	0,807	A	8	0,138	-0,532	5,500	2,563
					B	10	0,155	-0,218	8,889	2,619
					C*	38	0,690	0,630	13,725	3,552
					D	2	0,017	-0,224	4,000	0,000
2	38	0,603	0,568	0,810	A	9	0,155	-0,157	9,556	2,506
					B	5	0,086	-0,288	7,200	4,550
					C	9	0,155	-0,387	7,333	3,162
					D*	35	0,603	0,568	13,971	3,765
3	38	0,561	0,406	0,819	A*	32	0,561	0,406	13,656	3,832
					B	13	0,228	-0,230	9,308	5,023
					C	7	0,123	-0,286	7,857	2,545
					D	5	0,088	-0,067	10,200	4,382
4	38	0,690	0,630	0,807	A*	40	0,690	0,630	13,725	3,552
					B	9	0,155	-0,218	8,889	2,619
					C	8	0,138	-0,532	5,500	2,563
					D	1	0,017	-0,224	4,000	0,000
5	38	0,586	0,383	0,820	A*	34	0,586	0,383	13,441	3,661
					B	8	0,138	0,146	12,625	4,406
					C	3	0,052	-0,226	7,000	2,646
					D	13	0,224	-0,452	7,538	4,034
6	38	0,632	0,520	0,813	A	9	0,121	-0,013	11,143	2,610
					B	18	0,310	-0,047	11,000	3,896
					C*	20	0,379	0,410	14,500	3,726
					D	11	0,190	-0,442	7,455	4,525

Item	N	P	Total Rpbis	Alpha w/o	Option	N	Prop.	Rpbis	Mean	S
7	38	0,621	0,398	0,819	A	4	0,069	-0,310	6,250	6,076
					B	13	0,224	-0,144	9,923	3,639
					C*	36	0,621	0,398	13,361	4,044
					D	5	0,086	-0,193	8,400	2,966
8	57	0,614	0,383	0,820	A	2	0,034	-0,183	7,000	4,243
					B*	34	0,586	0,357	13,353	4,177
					C	14	0,241	-0,126	10,143	4,383
					D	8	0,138	-0,257	8,375	3,503
9*	38	0,586	0,357	0,821	A*	35	0,614	0,383	13,371	4,120
					B	8	0,140	-0,430	6,500	3,207
					C	11	0,193	0,096	11,909	2,773
					D	3	0,053	-0,352	4,667	0,577
10*	38	0,724	0,309	0,823	A	2	0,034	-0,131	8,000	2,828
					B	5	0,086	-0,270	7,200	5,119
					C*	42	0,724	0,309	12,762	4,119
					D	9	0,155	-0,106	9,889	4,781
11*	38	0,276	0,059	0,835	A	13	0,224	-0,181	9,923	4,645
					B	16	0,276	-0,012	11,313	4,332
					C	13	0,224	0,131	12,462	3,643
					D*	16	0,276	0,059	12,813	5,167

As seen in Table 3.8, 11 multiple-choice items were administered to 38 EFL students in pilot of EFL Achievement Test. Taking the stages of test construction explained in detail in the previous parts into consideration, the researcher prepared minimum 2 items to cover each of 5 objectives determined in the Table of Specifications for EFL Achievement Test (see Appendix 4). In this respect, 11 multiple items are matched with these 5 objectives as follows:

Table 3.9: Table of Specifications for the Piloted Multiple-Choice Item

Skill	Objectives & Functions	Understand	Apply	Analyze	Number of Items
Grammar	Students identify the correct form of the pronouns in a sentence.	X			*2* *(1,2)*
	Students use the correct quantifiers with countable and uncountable nouns.		X		*2* *(3, 11)*
	Students identify correct form of the adjectives to make comparison.	X			*2* *(4,9)*
	Students identify the correct tense to complete the sentences.	X			*3* *(5,6,8)*
	Students combine sentences by using correct conjunctions.		X		*2* *(7,10)*

As seen in Table 3.9, each objective that the students are expected to attain matches with minimum 2 multiple choice items. Taking the aforementioned acceptable ranges of item discrimination and difficulty and Turgut and Baykul's (2012) remark on item difficulty proportion of test items into account, the ultimate multiple-choice grammar section was formed.

In this respect, item-level statistics regarding 1st and 2nd items covering the 1st objective were reviewed. As seen in the table, the related items were found to have acceptable Rpbis > 0.30. On the other hand, these items were found to be moderately difficult with *P* between 0,35 – 0,64. With the aim of developing a moderately difficult test powerful enough to discriminate high scorers than low scorers, these two items were included in the ultimate test.

As seen in Table 3.8, 3rd item assessing the 2nd objective specified in Table 3.9 was explored to be a moderately difficult item with a *P* = 0,561 and an acceptable discriminative power with Rpbis > 0,30. However, the 11th item assessing the same objective was found to be difficult with a *P* in the range from 0,20 to 0,34. This item was excluded from the ultimate test since it was also found to have an unacceptable Rpbis of 0,059 indicating nearly no correlation between the item and the total scores.

As a result of reviewing item-level statistics for 4th and 9th items assessing the 3rd objective specified in Table 3.9, it was decided to exclude the 9th item keeping the 4th in the ultimate test. The reason behind this was that the 4th item differentiates high scorers than the low scorers with a Rpbis of 0,630 better than 9th item with a Rpbis of 0,357.

Review of statistics on the 5th, 6th and 8th items assessing 4th objective specified in Table 3.9 made the researcher keep these items in the ultimate test since all these 3 items were found to serve for the purpose of forming a moderately difficult test with *P* in the range from 0,35 to 0,64. These items were also explored to have acceptable Rpbis > 0,30.

Finally, item-level statistics on 7th and 10th items for the 5th objective revealed 7th item as a moderately difficult item with a Rpbis higher than that of 10th item which was to be easy with a *P* in the range from 0,65 to 0, 79.

As a result of these item-level statistics on 11 multiple-choice grammar items administered to 38 EFL students, 8 items, 7 moderately difficult and 1 easy, with Rpbis > 0,30 were included in the ultimate test. The ultimate table of specification for this part was formed as follows:

Table 3.10: Table of Specifications for the Ultimate Multiple-Choice Grammar Items

Skill	Objectives & Functions	Understand	Apply	Analyze	Number of Items
Grammar	Students identify the correct form of the pronouns in a sentence.	X			2 (1,2)
	Students use the correct quantifiers with countable and uncountable nouns.		X		1 (3)
	Students identify correct form of the adjectives to make comparison.	X			1 (4)
	Students identify the correct tense to complete the sentences.	X			3 (5,6,8)
	Students combine sentences by using correct conjunctions.		X		1 (7)

Following the item-level statistics on multiple choice items, the researcher took the step to calculate internal consistency of these items using KR_{20} since it was stated to

be a measure of internal consistency for tests or surveys which are composed of dichotomously scored items (Bademci, 2011; Tan, 2009). The following table presents the results for internal reliability.

Table 3.11: Reliability Statistics for the Ultimate Multiple-Choice Grammar Items

Score	Kr_{20}	SEM
Scored items	0,71	1, 25

As seen in the table, Kr_{20} of 0,71 indicates that this multiple choice part is a reliable assessor of EFL performance on the specified objectives. Standard error of measurement (SEM) stands for an index of the expected variation in observed scores of the test-takers due to measurement error. In other words, as set by Harvill (1991), SEM is calculated to interpret the difference in the obtained and true scores. On an achievement test, a SEM of $\pm$ 2 indicates that the researchers or teachers can be 95 % confident that true scores fall between 2 x SEM + observed scores and 2 x SEM – observed scores (Harvill, 1991). In this respect, as seen in the table, SEM of 1,25 indicates that this part is able to accurately estimate true scores of the test-takers and produces a confidence band close to true scores with 95 % accuracy.

In this part of the study, item-level statistics together with calculation of KR_{20} for the internal consistency of the second sub-section of grammar part were presented. This section included 8 "fill in the blank" items. These were organized as 4 different question and answer dialogues. Item-level statistics for this section are presented as follows:

Table 3.12: Item Statistics of the Piloted Fill in the Blank Grammar Items

Items	N	*P*	Rpbis
1	38	0,66	0,69
2	38	0,61	0,48
3	38	0,53	0,60
4	38	0,63	0,48
5	38	0,68	0,53
6	38	0,63	0,63
7	38	0,63	0,61
8	38	0,55	0,61

As seen in Table 3.12, 8 *fill in the blank* items in a limited response task were administered to 38 EFL students in the pilot of EFL Achievement Test. These items were aimed to measure the attainment of one objective determined in the Table of Specifications for EFL Achievement Test (see Appendix 4). In this respect, 8 items are matched with one objective as follows:

Table 3.13: Table of Specifications for the Piloted Fill in the Blank Grammar

Skill	Objectives & Functions	Understand	Apply	Analyze	Number of Items
Grammar	Students identify the tenses to complete the sentences with the correct form of the verbs.		X		*8* *(B.II.)*

Items

Taking the aforementioned acceptable ranges of item discrimination and difficulty into account, the ultimate fill in the blank grammar section was formed.

In this respect, item-level statistics regarding 8 items covering the objective specified in Table 3.13 were reviewed. As seen in the table, the related items were found to

have acceptable Rpbis > 0.30. On the other hand, 6 of these items were found to be moderately difficult with *P* between 0,35 – 0,64. Remaining 2 of them were explored to be easy with P in the range from 0,65 to 0,79. With the aim of developing a moderately difficult test with a high discriminative power (Rpbis > 0.30), all these items were included in the ultimate test. Following Turgut and Baykul's (2012) remark on item difficulty proportion of test that advises the test developers to include 13 % easy questions in their ultimate tests, 1st and 5th items with P in the range from 0,65 to 0,79 were not excluded from the test. In this respect, both of the items in the second sub-section in the grammar part of the test and the 6th objective they were planned to measure were kept in the ultimate test and table of specification without being exposed to any changes.

Table 3.14: Reliability Statistics for the Ultimate Fill in the Blank Grammar Items

Score	Kr_{20}	SEM
Scored items	0,72	1, 21

As seen in Table 3.14, Kuder-Richardson Formula 20 (Kr_{20}) was administered to 8 fill in the blank items measuring the 6th objective specified in table 3.13 to get an index of internal consistency of the section. Based on the acceptable ranges of KR_{20} for teacher-made tests (Cortina, 1993; McGahee, Ball, 2009; Sabri, 2013), this part was found to be reliable assessor of students' performance on the specified objective (KR_{20} = 0,72). Moreover, the table sets that the section measures EFL students' performance on the specified objective 2x1,21 + their observed scores and 2x1,21 – their observed scores with a confidence band close to true scores with 95 % accuracy (SEM = 1,24).

Following item-level statistics run for "fill in the blank" grammar section, the researcher analyzed the item difficulty and discrimination indices of 2 re-write items in addition to its index of internal consistency. This 3rd sub-grammar section included 2 items that require students use a grammatical sentence equivalent in meaning to the one that is given.

Table 3.15: Item Statistics of the Piloted Rewrite Items

Items	N	*P*	Rpbis
1	38	0,42	0,93
2	38	0,50	0,94

Table 3.15 presents the results of the item-level analysis of 2 re-write items administered to 38 EFL students in pilot of EFL Achievement Test. These items were aimed to measure the attainment of one objective determined in the Table of Specifications for EFL Achievement Test (see Appendix 4). In this respect, 2 items are matched with the following objective:

Table 3.16: Table of Specifications for the Piloted Rewrite Items

Skill	Objectives & Functions	Understand	Apply	Analyze	Number of Items
Grammar	Students rewrite the sentences by using the given structures and words.		X		*2(B.III)*

Taking the aforementioned acceptable ranges of item discrimination into account, the ultimate rewrite grammar section was formed.

In this respect, item-level statistics regarding 2 items covering the objective specified in Table 3.16 were reviewed. As seen in the table, the related items were found to have acceptable Rpbis > 0.30. In other words, these items were found to have a high discriminative power close to 1. On the other hand, these items were found to be moderately difficult with *P* between 0,35 – 0,64. The items were explored to have difficulty index close to the range from 0,20 to 0,34 indicating difficult items (*P*= 0,42; 50). However, this level of difficulty on these items was expected by the researcher since the task requires the test takers to rewrite, manipulate and use the given structures in a different sentence. Turgut and Baykul's (2012) remark on item difficulty proportion of test that advises the test developers to include 2 % of the

questions in their ultimate tests to be difficult. In addition to it, their high discriminative power drove the researcher to keep the items in this section and the 7th objective determined in the table of specifications without any changes for EFL Achievement Test (see Appendix 4).

Table 3.17: Reliability Statistics for the Ultimate Rewrite Items

Score	Kr_{20}	SEM
Scored items	0,85	0, 36

Table 3.17 presents the results of the analysis of Kuder-Richardson Formula 20 (Kr_{20}) on 2 rewrite items measuring the 7th objective specified in Table 3.16 to get an index of internal consistency of the section. Based on the acceptable ranges of KR_{20} for teacher-made tests (Cortina, 1993; McGahee, Ball, 2009; Sabri, 2013), this part was found to be highly reliable assessor of students' performance on the specified objective ($KR_{20} = 0,85$). On the other hand, SEM reveals the dispersion of the measurement errors. SEM of 0,36 indicates that this 3rd sub-section measures EFL students' true performance on the specified objective fluctuating from 2 x 0.36 + observed scores and 2 x 0,36 – observed scores with a 95 % confidence band.

As a result of item level analyses administered on three sub-sections of grammar part of the piloted EFL Achievement Test, ultimate test was decided to include 18 items, 8 multiple-choice, 8 fill in the blanks and 2 rewrite, measuring 7 critical objectives. The statistics for the ultimate grammar part of the achievement test was presented as follows:

Table 3.18: Summary Statistics for the Ultimate Grammar Section

Score	Items	Mean	S	Min Score	Max Score	Mean P	Mean Rpbis
Scored Items	18	9,97	4,78	0	17	0,55	0,54

Taking the acceptable discrimination and difficulty ranges covered in the related literature as the basis, the researcher came up with the 18-item grammar section. First of all, as seen in the table, there was a spread of scores, with a s = 4.78 and a range of 0 to 17. Mean of item difficulty in the grammar section indicates that items are moderately difficult with $P = 0.55$. On the other hand, item discrimination (Rpbis), indicates that this section differentiates between participants with high and low scores with a highly positive Rpbis at 0,54. As a result, the 18-item grammar section was found to have mean of item discrimination and difficulty within acceptable ranges.

Table 3.19: Reliability Statistics for the Ultimate Grammar Section

Score	Kr_{20}	SEM
Scored items	0,85	1, 83

As seen in the table, based on the acceptable ranges of KR_{20} for teacher-made tests (Cortina, 1993; McGahee, Ball, 2009; Sabri, 2013), Kr_{20} of 0,85 indicates that this grammar section is highly reliable assessor of EFL performance on the specified objectives. On the other hand, SEM of 1,83 indicates that the section interprets true scores of the test-takers in the range from + 2 x SEM and – 2 x SEM with a 95 % confidence band.

3.7.1.2. Pilot of Listening Part

In the listening section of the test, students' performance on 2 critical objectives was aimed to be evaluated. Due to item analyses that may result in removing some of the items due to their poor qualities, the researcher aimed to initially include at least two items to test each of the objectives in the table of specifications. As a result, 10 items in 2 different sub-sections were exposed to item-level statistics. While the first five items were designed as true-false items in a closed-ended task, the second five items were designed as completion items in a limited response task. The results of the analyses after piloting the test with 38 EFL students were presented in detail as follows:

Table 3.20: Summary Statistics for the Piloted Listening Section

Score	Items	Mean	S	Min Score	Max Score	Mean P	Mean Rpbis	KR$_{20}$	SEM
Scored Items	10	5,42	2,87	0	10	0,54	0,57	0,77	1,37

As seen in the table, listening scores of EFL students were spread with a s = 2.87 and a range of 0 to 10. Mean of items difficulty in the section indicates that items are moderately difficult with *P* in the range from 0,35 to 0,64 (P = 0.54). These items were also explored to have highly acceptable Rpbis > 0,30 differentiating between participants with high and low listening scores (Rpbis = 0,57).

In terms of the index of internal consistency of the section, based on the acceptable ranges ≥ 0,70, Kr$_{20}$ of 0,77 indicates that this piloted part, as a whole, is a reliable assessor of EFL listening performance on the specified objectives. Finally, a SEM of 1,37 reveals that the sections measure true listening scores of the test-takers in the range from + 2 x 1,37 to – 2 x 1,37 with 95 % accuracy.

The statistics presented above only provide information for the entire listening section. In this part there are two sub-sections assessing the participant students' acquisition of 2 different objectives. These objectives are assessed through 10 items. For quality control of each of these sub-sections, the researcher employed item-level analyses. The results of the analyses for the first part comprised of 5 true-false items were presented as follows:

Table 3.21: Item Statistics of the Piloted True-False Listening Items

Items	N	*P*	Rpbis
1	38	0,58	0,78
2	38	0,58	0,78
3	38	0,63	0,71
4	38	0,63	0,61
5	38	0,42	0,56

As seen in Table 3.21, 5 true-false items in a closed-ended task were administered to 38 EFL students in pilot of EFL Achievement Test. These items were aimed to measure the attainment of one objective determined in the Table of Specifications for EFL Achievement Test (see Appendix 4). In this respect, 5 items are matched with the following objective:

Table 3.22: Table of Specifications for the Piloted True-False Listening Items

Skill	Objectives & Functions	Understand	Apply	Analyze	Number of Items
Listening	Students distinguish true and false sentences in an audio text.			X	5 (A.I.)

Taking the aforementioned acceptable ranges of item discrimination and difficulty into account, the ultimate closed ended listening section was formed.

In this respect, item-level statistics regarding 5 items covering the objective specified in Table 3.22 were reviewed. As seen in the table, the related items were found to have highly acceptable Rpbis > 0.30. On the other hand, each of these items were found to be moderately difficult with P between 0,35 – 0,64. With the aim of developing a moderately difficult test with a high discriminative power (Rpbis > 0.30; P = 0,35 – 0,64), all these items were included in the ultimate test. In this respect, no edition was administered on neither the individual items in this sub-section nor the related part in table of specification.

Table 3.23: Reliability Statistics for the Ultimate True-False Listening Items

Score	Kr_{20}	SEM
Scored items	0,72	0, 90

As seen in Table 3.23, Kuder-Richardson Formula 20 (Kr_{20}) was administered to 5 true-false listening items measuring the 1st listening objective specified in Table 3.22

to get an index of internal consistency of the section. Based on the acceptable ranges of KR_{20} for teacher-made tests (Cortina, 1993; McGahee, Ball, 2009; Sabri, 2013), this part was found to be reliable assessor of students' listening performance on the specified objective ($KR_{20} = 0,72$). Moreover, the table sets that the section measures EFL students' true listening performance closely to their observed scores with a 95 % confidence band (SEM = 0,90).

Following item-level statistics run for the closed-ended listening section, the researcher computed the item difficulty and discrimination indices of 5 listening completion items in addition to its index of internal consistency. This section was designed as a limited response task comprised of 5 items that require students to recognize, recall, and select the relevant information on five different conversations they hear to complete the guided short answers. The results of item level analyses on the section were shown in the table below as follows:

Table 3.24: Item Statistics of the Piloted Listening Completion Items

Items	N	*P*	Rpbis
1	38	0,53	0,76
2	38	0,53	0,60
3	38	0,50	0,71
4	38	0,58	0,67
5	38	0,45	0,66

As seen in Table 3.24, 5 limited response items were administered to 38 EFL students in the pilot of EFL Achievement Test to measure the attainment of one objective determined inthe Table of Specifications for EFL Achievement Test (see Appendix 4). The related part of the Table of Specifications was presented as follows:

Table 3.25: Table of Specifications for the Piloted Listening Completion Items

Skill	Objectives & Functions	Understand	Apply	Analyze	Number of Items
Listening	Students complete the blanks in an audio text.	X			*5* *(A.II.)*

Taking the aforementioned acceptable ranges of item difficulty and discrimination into account, the researcher decided on the ultimate limited response listening section. Reviewing 5 items covering the objective specified in Table 3.25, the researcher decided to keep all the items in the ultimate test without a change. The reason behind this was that, the related items were found to have highly acceptable Rpbis > 0.30 in a moderately difficult range with *P* between 0,35 – 0,64. Although the 5th item was found to be more difficult than the others, it was needed to keep some difficult items in the ultimate test.

Table 3.26: Reliability Statistics for the Ultimate True-False Listening Items

Score	Kr_{20}	SEM
Scored items	0,71	0, 92

As seen in Table 3.26, Kuder-Richardson Formula 20 (Kr_{20}) was administered to 5 listening completion items measuring the 2nd listening objective specified in table 3.25 to get an index of internal consistency of the section. Based on the acceptable ranges of KR_{20} for teacher-made tests (Cortina, 1993; McGahee, Ball, 2009; Sabri, 2013), this part was found to be reliable assessor of students' listening performance on the specified objective measuring EFL students' true listening performance very closely to their observed scores with a 95 % confidence band (KR_{20} = 0,71; SEM = 0,92).

In conclusion, taking the acceptable discrimination and difficulty ranges covered in the related literature as the basis, the researcher came up with the 10-item listening

section organized in 2 sub-sections. Since all the items produced indices in the acceptable ranges, no edition was employed on the items by the researcher. In this respect, as specified in Tables 3.21 and 3.24, the listening part was found to be a moderately difficult test with the mean of P in the range 0,35 to 0,64 consistently discriminating between high scorers and low scorers (Mean of Rpbis = 0, 57; KR_{20} = 0, 77).

3.7.1.3. Pilot of Reading Part

In the reading section of the achievement test, students' performance on 3 critical objectives was aimed to be evaluated. Due to item level analyses that may result in removing some of the poor items, the researcher aimed to initially include at least two items to test each of the objectives in the table of specifications. As a result, 15 items in 3 different sub-sections about a reading passage comprised of approximately 350 words in four paragraphs were exposed to item-level statistics. While the first five items were designed as true-false items in a closed-ended task, the second five items were designed as another closed ended task requiring the students to find whether some specific information exists in the passage. On the other hand, the third five items were organized as a limited response task where the students are required to scan the text to answer the questions about the passage. The results of the analyses after piloting the test with 38 EFL students were presented in detail as follows:

Table 3.27: Summary Statistics for the Piloted Reading Section

Score	Items	Mean	S	Min Score	Max Score	Mean P	Mean Rpbis	KR_{20}	SEM
Scored Items	15	6,97	4,66	1	15	0,46	0,63	0,89	1,56

As seen in the table, reading scores of EFL students were spread with a s = 4.66 and a range of 1 to 15. On the other hand, mean of item difficulty in the section indicates that items are moderately difficult with P in the range from 0,35 to 0,64 (P = 0.46). while these items highly acceptable Rpbis values > 0,30 differentiating between participants with high and low reading scores (Rpbis = 0,63).

Based on the acceptable ranges of $KR_{20} \geq 0,70$ for teacher-made tests, the table indicates that this piloted reading part, as a whole, is a reliable assessor of EFL reading performance on the specified objectives with a SEM measuring true scores of the participants in the range from + 2 x 1,56 to – 2 x 1,56 with 95 % accuracy. ($KR_{20} = 0,89$; SEM = 1,56).

The statistics presented above only provide information for the entire reading section. In this part there are three sub-sections assessing the participant students' acquisition of 3 different objectives. These objectives are assessed through 15 items. For quality control of each of these sub-sections, the researcher employed item-level analyses. The results of the analyses for the first part comprised of 5 true-false items were presented as follows:

Table 3.28: Item Statistics of the Piloted True-False Reading Items

Items	N	*P*	Rpbis
1	38	0,53	0,68
2	38	0,53	0,80
3	38	0,47	0,74
4	38	0,32	0,67
5	38	0,55	0,58

As seen in Table 3.28, 5 true-false items in a closed-ended task were administered to 38 EFL students in the pilot of EFL Achievement Test. These items required the students to distinguish the sentences and identify them as true or false (see Appendix 4). In this respect, 5 items are matched with the following objective:

Table 3.29: Table of Specifications for the Piloted True-False Reading Items

Skill	Objectives & Functions	Understand	Apply	Analyze	Number of Items
Reading	Students distinguish true and false sentences in a text.			X	*5 (C.I.)*

Taking the aforementioned acceptable ranges of item discrimination and difficulty into account, the ultimate closed ended section was formed.

In this respect, item-level statistics regarding 5 items covering the objective specified in Table 3.29 were reviewed. As seen in the table, the related items were found to have highly acceptable Rpbis > 0.30. On the other hand, four of these items were found to be moderately difficult with P between 0,35 – 0,64. However, the 4th item was seen to be a difficult one in the range from 0,20 to 0,34 with a high discriminative power (Rpbis= 0,67). Taking these difficulty ranges and Turgut and Baykul's (2012) remark on item difficulty proportion of test items into account, 13 % of the items were supposed to be difficult. In this respect, all the items in this part were decided to be kept in the ultimate test without employing a change.

Table 3.30: Reliability Statistics for the Ultimate True-False Reading Items

Score	Kr_{20}	SEM
Scored items	0,73	0,90

As seen in Table 3.30, Kuder-Richardson Formula 20 (Kr_{20}) was administered to 5 true-false reading items measuring the 1st reading objective specified in Table 3.29 to get an index of internal consistency of the section. Based on the acceptable ranges of KR_{20} for teacher-made tests (Cortina, 1993; McGahee, Ball, 2009; Sabri, 2013), this part was found to be a reliable assessor of students' reading performance on the specified objective ($KR_{20} = 0,73$). Moreover, the table sets that the section measures

EFL students' true reading performance closely to their observed scores ± 2 SEM with a 95 % confidence band (SEM = 0,90).

Following item-level statistics on the closed-ended true-false reading section, the researcher analyzed the item difficulty and discrimination indices of 5 reading items designed as another closed-ended task where the students were asked to scan the text to find out whether some specific information exists in the text or not. The results of item level analyses on the section were shown in the table below as follows:

Table 3.31: Item Statistics of the Piloted Closed-Ended Reading Items

Items	N	*P*	Rpbis
1	38	0,53	0,63
2	38	0,34	0,72
3	38	0,53	0,57
4	38	0,50	0,71
5	38	0,50	0,77

As seen in Table 3.31, 5 true-false items in a closed-ended task were administered to 38 EFL students in the pilot of EFL Achievement Test. These 5 items are matched with the following objective:

Table 3.32: Table of Specifications for the Piloted Closed–Ended Reading Items

Skill	Objectives & Functions	Understand	Apply	Analyze	Number of Items
Reading	Students make inferences from a text by using the given clues.			X	*5* *(C.II.)*

Taking the aforementioned acceptable ranges of item discrimination and difficulty into account, the ultimate closed ended reading section was formed.

In this respect, item-level statistics regarding 5 items covering the objective specified in Table 3.32 were reviewed. As a result, all the items in this part were decided to be kept in the ultimate test without employing a change. The reason behind this decision was the acceptable item discrimination and difficulty indices of the mentioned items (Rpbis > 0.30). On the other hand, four of these items were found to be moderately difficult with P between 0,35 – 0,64. However, 2nd item in this part was explored to be a difficult one in the range from 0,20 to 0,34 with a high discriminative power (Rpbis= 0,72). Thinking that keeping this item in the ultimate test would serve for the ideal item difficulty proportion of test items, the researcher included this item in the finalized test.

Table 3.33: Reliability Statistics for the Ultimate Closed-Ended Reading Items

Score	Kr_{20}	SEM
Scored items	0,71	0,92

As seen in Table 3.33, Kuder-Richardson Formula 20 (Kr_{20}) was administered to 5 closed ended items measuring the 2nd reading objective specified in table 3.32 to get an index of internal consistency of the section. As a result, KR_{20} of 0,71 revealed this part as a reliable assessor of students' reading performance on the specified objective measuring their true reading performance closely to their observed scores $\pm$ 2 SEM with a 95 % confidence band (SEM = 0,92).

Following item-level statistics on the closed-ended reading sections, the researcher took the step to compute the item difficulty and discrimination indices of 5 limited response reading items. These items required the students to comprehend the text and answer the questions about it. The results of item level analyses on the items in the section were shown in the table below:

Table 3.34: Item Statistics of the Piloted Limited Response Reading Items

Items	N	*P*	Rpbis
1	38	0,34	0,65
2	38	0,32	0,79
3	38	0,47	0,85
4	38	0,47	0,76
5	38	0,58	0,75

As seen in Table 3.34, 5 limited response reading items were administered to 38 EFL students in the pilot of EFL Achievement Test. These items required the students to scan the text to answer the related 5 questions. In this respect, 5 items are matched with the following objective:

Table 3.35: Table of Specifications for the Piloted Limited Response Reading Items

Skill	Objectives & Functions	Understand	Apply	Analyze	Number of Items
Reading	Students answer wh- questions about a given text.	X			*5* *(C.III)*

Taking the aforementioned acceptable ranges of item discrimination and difficulty into account, the researcher decided to keep all the items in the ultimate test. As seen in the table, the related items were found to have highly acceptable Rpbis > 0.30. On the other hand, three of these items were found to be moderately difficult with *P* between 0,35 – 0,64. However, the 1st and 2nd items were seen to be difficult having difficulty indices in the range from 0,20 to 0,34. Taking Turgut and Baykul's (2012) remark on item difficulty proportion of test items and enforcing nature of the items in

this section into account, these two difficult items were included in the ultimate test thanks to their high discriminative powers (Rpbis = 0,65; 0,79). In this respect, all the items in this were decided to keep in the ultimate test without employing a change.

Table 3.36: Reliability Statistics for the Ultimate Limited Response Reading Items

Score	Kr_{20}	SEM
Scored items	0,82	0,80

As seen in Table 3.36, Kuder-Richardson Formula 20 (Kr_{20}) was administered to 5 limited response reading items measuring the 3rd reading objective specified in Table 3.35 to get an index of internal consistency of the section. Based on the acceptable ranges of KR_{20} for teacher-made tests (Cortina, 1993; McGahee, Ball, 2009; Sabri, 2013), this part was found to be a reliable assessor of students' reading performance on the specified objective (KR_{20} = 0,82). Moreover, the table sets that the section measures EFL students' true reading performance closely to their observed scores $\pm$ 2 SEM with a 95 % confidence band (SEM = 0,80).

To conclude, taking the acceptable discrimination and difficulty ranges and ideal population of the difficult items in the ultimate test into consideration, the researcher decided to preserve the ultimate reading part with 15 items in three sub-sections. Based on these results, the reading part was found to be a moderately difficult test with the mean of *P* in the range 0,35 to 0,64 consistently discriminating between high scorers and low scorers (Mean of Rpbis = 0, 63; KR_{20} = 0, 89).

3.7.1.4. Pilot of Vocabulary Part

In the vocabulary section of the test, students' performance on 2 critical objectives was measured. Due to item level analyses that may cause some of the poor items to be removed from the ultimate test, the researcher aimed to include at least two items to test each of the objectives in the table of specifications. While the first part was designed as 8 fill in the blank items in a closed-ended task, the second part was designed as 6 word-formation items in a limited response task. The results of the

analyses after piloting the test with 38 EFL students were presented in detail as follows:

Table 3.37: Summary Statistics for the Piloted Vocabulary Section

Score	Items	Mean	S	Min Score	Max Score	Mean P	Mean Rpbis	KR_{20}	SEM
Scored Items	14	6,21	3,76	0	14	0,44	0,54	0,81	1,64

As seen in the table, vocabulary scores of EFL students were spread with a s = 3, 76 and a range of 0 to 14. On the other hand, mean of item difficulty in the section were found to be moderately difficult with *P* in the range from 0,35 to 0,64 (P = 0.44) while these items have highly acceptable Rpbis values > 0,30 differentiating between participants with high and low vocabulary scores (Rpbis = 0,54).

Based on the acceptable ranges of $KR_{20} \geq 0,70$ for teacher-made tests, the table indicates this piloted part as a reliable assessor of EFL vocabulary performance on the specified objectives with a SEM measuring true scores of the participants 2 x SEM ± their observed scores with 95 % accuracy ($KR_{20} = 0,81$; SEM = 1,64).

In addition to the statistics presented above for the entire vocabulary section, the researcher employed item-level statistics for two sub-sections comprised of 14 items assessing the participant students' acquisition of 2 different objectives. The results of the analyses for the first part comprised of 8 fill in the blank items were presented as follows:

Table 3.38: Item Statistics of the Piloted Fill in the Blank Vocabulary Items

Items	N	*P*	Rpbis
1	38	0,45	0,68
2	38	0,53	0,67
3	38	0,50	0,69
4	38	0,50	0,73
5	38	0,50	0,52
6	38	0,47	0,34
7	38	0,39	0,62
8	38	0,45	0,68

As seen in Table 3.38, 8 fill in the blank vocabulary items in a closed-ended task were administered to 38 EFL students in pilot of EFL Achievement Test. These items were aimed to measure the attainment of one objective determined in the Table of Specifications for EFL Achievement Test where students were required to understand the meaning of the both sentences and words to complete the sentences with the correct word from the box (see Appendix 4). In this respect, 8 items are matched with the following objective:

Table 3.39: Table of Specifications for the Piloted Closed-Ended Vocabulary Items

Skill	Objectives & Functions	Understand	Apply	Analyze	Number of Items
Vocabulary	Students identify the best vocabulary item to complete a sentence.	X			*8* *(D.I.)*

Taking the aforementioned acceptable ranges of items discrimination and difficulty into account, the ultimate closed ended vocabulary section was formed.

In this respect, item-level statistics regarding 8 items covering the objective specified in Table 3.39 were reviewed. As seen in the table, the related items were found to have highly acceptable Rpbis > 0.30. On the other hand, each of these items were found to be moderately difficult with P between 0,35 – 0,64. With the aim of developing a moderately difficult test with a high discriminative power (Rpbis > 0.30; P = 0,35 – 0,64), all these items were included in the ultimate test. In this respect, no edition was administered on neither the individual items in this sub-section nor the related part in table of specification.

Table 3.40: Reliability Statistics for the Ultimate Closed-Ended Vocabulary Items

Score	Kr_{20}	SEM
Scored items	0,77	1,20

As seen in Table 3.40, Kuder-Richardson Formula 20 (Kr_{20}) was administered to 8 fill in the blank items measuring the 1st vocabulary objective specified in Table 3.39 to get an index of internal consistency of the section. Based on the acceptable ranges of KR_{20} for teacher-made tests (Cortina, 1993; McGahee, Ball, 2009; Sabri, 2013), this part was found to be a reliable assessor of students' vocabulary performance on the specified objective (KR_{20} = 0,77). Moreover, the table sets that the section measures EFL students' true vocabulary performance closely 2 x SEM ± their observed scores with a 95 % confidence band (SEM = 1,20).

Following item-level statistics on the closed-ended vocabulary sections, the researcher took the next step to employ the item-level on 6 limited response vocabulary items. These items require the students to choose the right words for each sentence and use the words in the correct form which suits the sentence best. The results of item level analyses on the items in the section were shown in the table below:

Table 3.41: Item Statistics of the Piloted Limited Response Vocabulary Items

Items	N	*P*	Rpbis
1	38	0,47	0,60
2	38	0,42	0,69
3	38	0,32	0,63
4	38	0,39	0,65
5	38	0,39	0,62
6	38	0,42	0,69

As seen in Table 3.41, 6 limited response vocabulary items task were administered to 38 EFL students in the piloted EFL Achievement Test. These items required the EFL students to modify the words given with the sentences to best complete the sentences (see Appendix 4). In this respect, 6 items are matched with the following objective:

Table 3.42: Table of Specifications for the Piloted Limited Response Vocabulary Items

Skill	Objectives & Functions	Understand	Apply	Analyze	Number of Items
Vocabulary	Students modify the words into correct form to complete the sentences.		X		*6* *(D.II.)*

Taking the aforementioned acceptable ranges of item discrimination and difficulty into account, the researcher decided to keep all the items in the ultimate test. As seen in the table, the related items were found to have highly acceptable Rpbis > 0.30. On the other hand, five of these items were found to be moderately difficult with P between 0,35 – 0,64. However, 3rd item was seen to be difficult having difficulty index in the range from 0,20 to 0,34. Taking Turgut and Baykul's (2012) remark on

item difficulty proportion of test items into account, this difficult item was included in the ultimate test thanks to its high discriminative powers (Rpbis = 0,63). In this respect, all the items in this part were decided to be kept in the ultimate test without employing a change.

Table 3.43: Reliability Statistics for the Ultimate Limited Response Vocabulary Items

Score	Kr_{20}	SEM
Scored items	0,72	1,02

As seen in Table 3.43, Kuder-Richardson Formula 20 (Kr_{20}) was administered to 6 word-formation items measuring the 2nd vocabulary objective specified in Table 3.42 to get an index of internal consistency of the section. Based on the acceptable ranges of KR_{20} for teacher-made tests (Cortina, 1993; McGahee, Ball, 2009; Sabri, 2013), this part was found to be a reliable assessor of students' vocabulary performance on the specified objective (KR_{20} = 0,72). Moreover, the table sets that the section closely measures EFL students' true vocabulary performance 2 x SEM ± their observed scores with a 95 % confidence band (SEM = 1,02).

To conclude, taking the acceptable discrimination and difficulty ranges and ideal population of the difficult items in the ultimate test into consideration, the researcher decided to preserve the ultimate vocabulary part with 10 items in two sub-sections as they were in the piloted test. Based on the results specified in Tables 3.38 and 3.41, the vocabulary part was found to be a moderately difficult test with the mean of *P* in the range 0,35 to 0,64 consistently discriminating between high scorers and low scorers (Mean of Rpbis = 0, 54; KR_{20} = 0, 81).

3.7.1.5. Pilot of Total EFL Achievement Test

In four separate parts above, the item and reliability analyses regarding listening, grammar, reading and vocabulary sections were covered. However, due to its open ended and creative nature, the writing part was excluded and assessed by three EFL teachers following the guidelines in the analytic rubric provided in the detailed answer key (see Appendix 2). Writing performances were assessed by following

guidelines for four main areas (grammar, vocabulary, content and organization) analytically and separately out of 5. Participants' EFL writing performance was ultimately produced capturing the mean of three EFL assessors. As a result, writing part was not exposed to item-level or reliability analyses. However, reliability was assessed through the mean scores of three assessors based on analytic rubric while validity was ensured consulting 5 EFL experts' views on the harmony between the objective and the item regarding writing.

In this respect, basic statistics for the total of EFL Achievement Test excluding writing part was presented in the table below as follows:

Table 3.44: Summary Statistics for the Piloted Total EFL Achievement Test

Score	Items	Mean	S	Min Score	Max Score	Mean P	Mean Rpbis	KR$_{20}$	SEM
Scored Items	57	28,58	10,65	9	52	0,51	0,68	0,91	3,20

As seen in the table, the researcher employed analyses for the ultimate entire EFL Achievement Test excluding the poor items and writing section. In this respect, total scores of EFL students in grammar, reading, listening and vocabulary were spread with a s = 10, 65 and a range from 0 to 14. Mean of a test is expected to be half of the number of the items (Dilbaz, Özgelen, Yelken, 2012). Therefore, mean scores of the participant was found to be 28.58 and this is psychometrically appropriate for a 57-item test ($\bar{x}$=25,58). On the other hand, mean of item difficulty of the test was found to be moderately difficult with *P*in the range from 0,35 to 0,64 (P = 0.51) while these items were explored to have highly acceptable Rpbis values > 0,30 differentiating between high achievers and low achievers (Rpbis = 0,68).

Based on the acceptable ranges of KR$_{20}$ ≥ 0,70 for teacher-made tests (Cortina, 1993; McGahee, Ball, 2009; Sabri, 2013), this piloted EFL test was indicated as a highly reliable assessor of EFL performance on the specified objectives with a SEM closely measuring true scores of the participants 2 x SEM ± their observed scores with 95 % accuracy (KR$_{20}$ = 0,91; SEM = 3,20).

In addition to item and reliability analyses followed during the process of constructing the achievement test, the researcher took the last step to validate the test as an assessor of progress the participants make during the English course in a prep EFL class at Gebze Technical University. In this respect, the researcher relied on criterion-related validity. This type of validity is adopted by comparing the results of a new measurement with another well-established measurement that acts as a criterion (Turgut, Baykul, 2012). As set bay Balcı (2005), these criterions are sometimes used as scores obtained from well-established scales or tests while they are sometimes used as groups who are thought to definitely have or lack of the construct measured by the validated scale or test. In this respect, the researcher adopted the scores of 35 freshmen students who were qualified to study at their departments after their one-year prep education at foreign language department at Gebze Technical University. These students were identified with the permission of the department and the students were invited through e-mails to take the exam. 35 students who passed proficiency exam in 2015 accepted to take the achievement test. Their scores were compared with the scores of 38 students who were qualified to study at English Prep Schools of eight different universities and participated in the pilot study. In this respect, to decide on the type of the data, *Kolmogorov Smirnov Test* was computed and it was found the data for the variable did not differ significantly from the normal distribution (z =,767; p >.05). Within this respect, in the final step of the related analysis, the researcher computed Independent Samples T-test, a parametric test, to see whether two groups differ significantly in terms of their scores of EFL achievement test. The results were shown in the following table.

Table 3.45: Independent Samples T-Tests for Validity of EFL Achievement Test

Factor	Groups	N	$\overline{\text{x}}$	S	$Sh_{\overline{\text{x}}}$	t	Df	p	d
EFL Achievement	Freshmen	35	71,11	14,78	2,49	7,360	71	,000	0,68
	Prep	38	45,97	14,38	2,33				

As shown in Table 3.45, Independent Samples T-Test was run to find out if there was a statistically significant difference between freshmen and prep students' mean

scores of ultimate EFL achievement test. The results indicated a significant difference (t = 7.360; p < .001). In other words, freshmen students who were exposed to prep instruction and pass the proficiency test in 2015 were found to be significantly more successful than those who could not pass the test and had to attend prep classes in 2016. Cohen's *d* (degree of impact) indicates the standardized difference between means for t-tests results (Cohen, 1994). Reporting the effect size of the significant difference between two means, "*d*" makes detailed analysis possible. In other words, regardless of the variables, sample sizes or measurement types, objective measure of the significant difference are provided. Although the interpretation of the effect of an intervention is not straightforward, based on Cohen's (1988) framework, three-folded guideline are usually referred as "small, *d* = .2," "medium, *d* = .5," and "large, *d* = .8" (Ruscio, 2008). Within this respect, based on ranges suggested by Cohen himself *Cohen's d* of 0.68 suggested moderate practical significance in the favour of freshmen students (*d* > .50; *d* < .80). As a result of this analysis, taking the freshmen students' performance in EFL achievement test as a reference to make a comparison, the ultimate test was found to be a valid assessor of the prep students' EFL performance at Gebze Technical University.

3.7.2. Pilot of EFL Self-Efficacy Belief Scale

To explore the impact of flipped classroom teaching model on EFL learners' self-reported efficacy beliefs in EFL, the researcher required to employ a self-efficacy scale as a pre- and post-test. Therefore, *Self-Efficacy English Scale* developed by Yanar and Bümen (2012) to measure high school EFL learners' self-reported efficacy beliefs in four main ability domains of English as listening, reading, speaking and writing was piloted with 38 students. These students were qualified to study at English Prep Schools of eight different universities (5 state and 3 private) in Turkey in 2016-2017. The main purpose of employing this scale in the pilot study was to test its statistical relevance to the sample of the study. In other words, following the procedures to take the permission to use the scale for the current study, the researcher employed *LISREL 8.54 for Windows* based on a full structural equation model (SEM) to test and confirm the four-dimensioned model of the adopted scale in the pilot study.

Within this perspective, analysis of the data gathered from the pilot study was four-folded. In the first fold, Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) & Bartlett's Test of Sphericity was computed to measure the adequacy of the sampling to administer intended analyses. In the second fold, the researcher computed the Kolmogorov-Smirnov test to assess the normality of the data. This test is statistically a prerequisite for deciding on the best estimation method to adopt in CFA (Çapık, 2014; Şimşek, 2007). In the following step, the researcher administered confirmatory factor analyses (CFA) to explore if the scale is a valid measure of the construct "self-efficacy beliefs in EFL", which was explored and confirmed as a model by Yanar and Bümen (2012) to explain 61,41 % of the total variance with α = 0.97. Within this, suitable estimation method, fit indices within acceptable ranges specified before (Harrington, 2008; Schumacker, Lomax, 2004; Şimşek, 2007) and path diagrams were taken into account to evaluate the results related to the model. In the fourth fold, reliability coefficient analysis to elicit cronbach's alpha (α) value through *SPSS 21 for Windows* was run.

The first fold was devoted to measuring the adequacy of the sampling in the pilot testing before checking the reliability and confirming the construct validity of the instrument. Therefore, *KMO & Bartlett's Test of Sphericity* was computed through *SPSS 21 for Windows*. As a result, the measure of the sampling adequacy was found to be higher than minimum acceptable value set by Kaiser (1960) as greater than 0.5. This KMO and Bartlett's significance < 0.001 show that the sample size in the pilot study is sufficient to conduct the intended factor analysis (KMO = 0.55; x^2 = 1214.138; p < .001).

Once the sample was found to be suitable to conduct intended analyses, the researcher set to test the normality of the distribution by employing the *Kolmogorov-Smirnov Test t*hrough *SPSS 21 for Windows* to decide on the best estimation method in CFA. The results of this test of normality were presented in the following table:

Table 3.46: Kolmogorov-Smirnov Test for the Normality of the Scores of EFL Self-Efficacy Beliefs Scale

		Self-Efficacy Beliefs
N		38
Parameters	$\overline{x}$	2,7646
	S	,59498
K-Smirnov Z		,616
P		,843

To come up with a decision whether the data distributes normally or not, *Kolmogorov Smirnov test* was run on the data gathered from the administration of EFL Self-Efficacy Scale in the pilot study. As a result, as seen in the table, the data for the variable was found not to differ significantly from the normal distribution (z=,616; p >.05). In this respect, Maximum Likelihood Method was chosen as the estimation method in CFA. Maximum Likelihood Estimation Method is one of the best and most frequently employed estimation method in CFA when the data is made of continuous variables that distribute normally (Şimşek, 2007).In the third fold, the researcher aimed to test the statistical relevance of the adopted scale through CFA based on Maximum Likelihood Method using *LISREL 8.54 for Windows*. In this respect, the researcher attempted to confirm how much of four latent variables (self-efficacy beliefs in EFL reading, writing, listening and speaking) were explained by 34 observed variables. Latent variables were shown in ellipses while 34 observed variables were shown in 34 rectangles. The factor loadings of observed variables from Item1 to Item34 on latent variables were indicated in the following two figures:

Chi-Square=652.57, df=523, P-value=0.00009, RMSEA=0.082
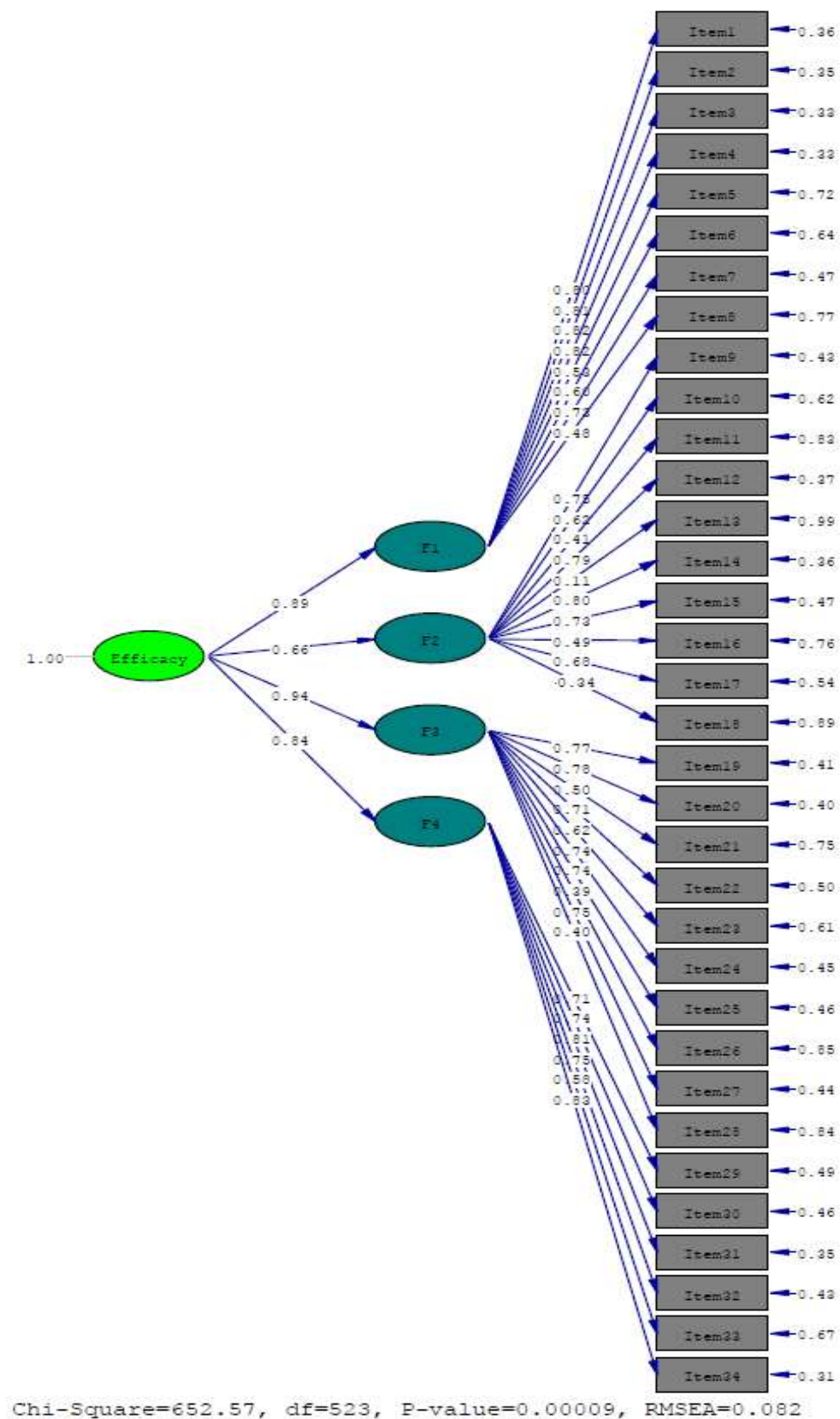
Figure 3.11: Model and Standardized Solutions for EFL Self-Efficacy Belief Scale

In the Figure 3.11, the standardized solution results indicate that the observed 34 variables represent the latent variable (self-efficacy beliefs) in four sub-scales ranging from 0.11 to 0.83. These results confirm the four-sub-scaled model set by Yanar and Bümen (2012). In other words, standardized solutions produce similar results to the explored model where 34 items were found together in four factors with extraction values ranging 0.42 to 0.69. As seen in the figure, with a perfect match with the explored model, 8 observed variables were found to explain F1 (reading sub-factor) while 10 observed variables were confirmed to explain F2 (writing sub-factor). On the other hand, 10 variables from 19 to 28 were found to explain F3 (listening sub-factor) while the last 6 variables were found to explain F4 (speaking sub-factor).

Factor loadings of an observed variable is expected to be no less than 0.30 and error covariance values have to be low to better explain a latent variable (Harrington, 2008; Şimşek, 2007). In this respect, all the factor loadings excluding that of 13th item were found to be acceptable ranging from 0,34 to 0,83. On the other hand, error covariance values except for those of the 13th and 18th items were found to well explain the latent variable in acceptable ranges of error. When the error covariance values and factor loadings values are analyzed together, it is seen that all variables except for 13th and 18th observed variables with an error covariance of 0,99; 089 and a factor loading of 0,11; 0,34 respectively were found to be in acceptable ranges. However, the perfect match between the explored four sub-scaled model and the model confirmed by the researcher indicate four sub-factors explain the main latent variable (EFL self-efficacy belief) with factor loadings > 0,63 and 0,71 standing for very good and perfect ranges (Harrington, 2008). In addition, there are some other reasons of adding this item together with its error into the model. First one of them is the confirmed four-dimension construct of the adopted scale. Secondly, the views of the consulted specialists that the items represent and explain the self-efficacy beliefs in EFL writing sufficiently make it difficult to extract the item. Based on these explained facts, all these 34 variables were kept and added into the model as explored by Yanar and Bümen (2012).

Chi-Square=652.57, df=523, P-value=0.00009, RMSEA=0.082
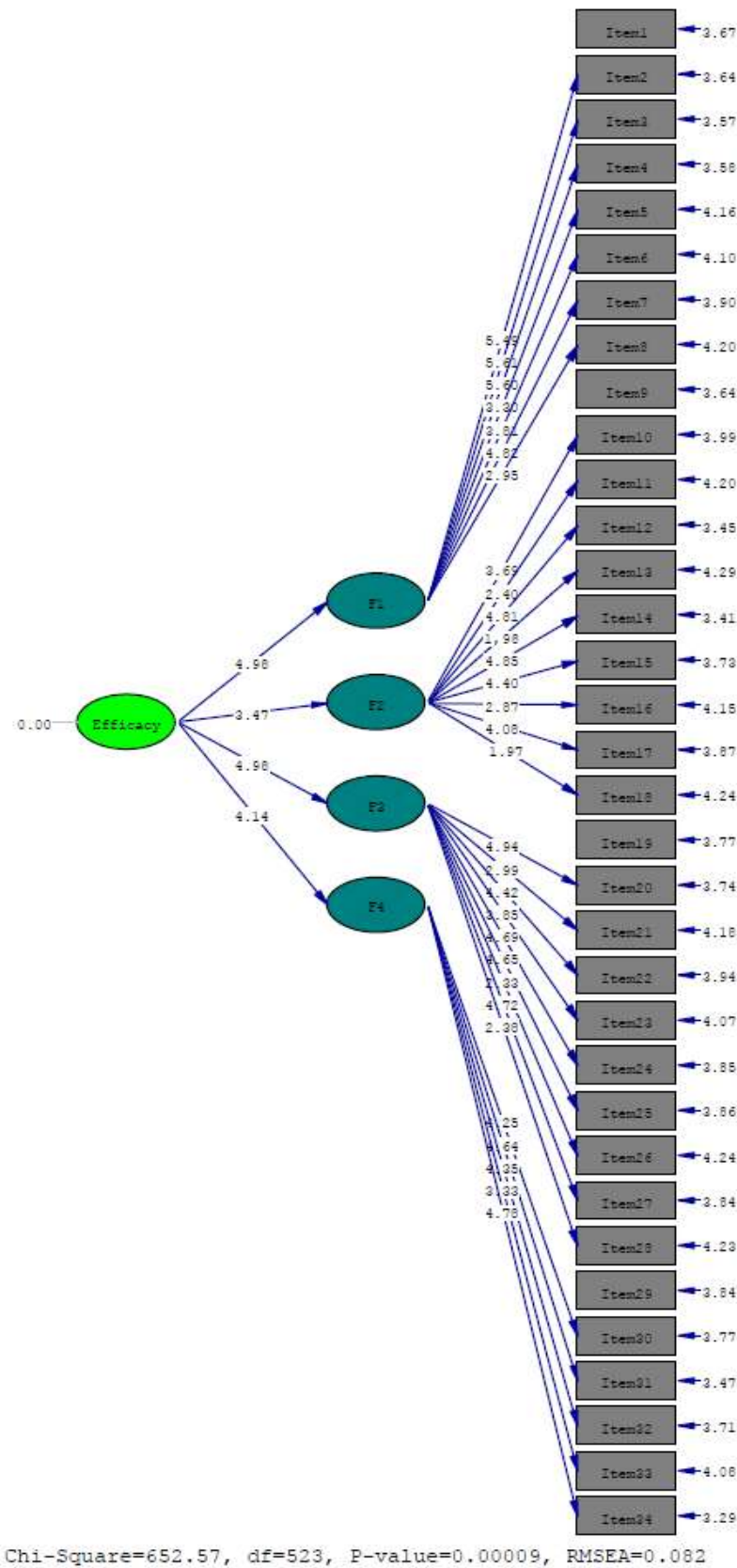
Figure 3.12: Not Standardized T-Values for EFL Self-Efficacy Belief Scale

T-values in Figure 3.12. refer to another parameter confirming and exploring the relationships between the observed and the latent variables. The figure shows that the t-values range from 1.97 to 5.61. Statistically, when the significance. 05 is accepted as a criterion in social sciences, t-values must be 1.96 or above it while it must be 2.56 or above it at 0,001 significance level (Şimşek, 2007; Çapık, 2014). Therefore, the variables having t-value less than 1.96 are advised to be omitted from the scale. However, no variable, as seen in the figure, has a t-value below 1.96. Moreover, 25 of the observed variables have t-values greater than 2.56 ($p < 0.001$) while 9 of them have t-values in the range from 1.96 to 2.55 ($p < 0.05$). As a result of these not standardized t-values, the adopted scale is confirmed to be a valid predictor of university prep students' EFL self-efficacy beliefs.

In addition to standardized solutions and t-values obtained from path diagrams as shown in the figures above, the researcher also checked fit indices to confirm the scale as a valid assessor of EFL self-efficacy beliefs in four dimensions. Therefore, following table presents the necessary fit indices related to the model.

Table 3.47: Fit Indices for EFL Self-Efficacy Belief Scale

Fit Indices	Values
x^2	652.27
Df	523
Chi-square (x^2) / df	1,24
GFI	0.79
AGFI	0.82
CFI	0.87
RMSEA	0.082
SRMR	0,013
RMR	0,013
NFI	0.85
IFI	0.88

Fit indices were taken into account to evaluate the results related to the model since they serve to confirm if the explored model fits with the data or the variables

(Bryant, Yarnold, 1995). To make a conclusive remark on the scale, the values in Table 3.47 were evaluated following the acceptable ranges specified in the literature (Harrington, 2008; Schumacker, Lomax, 2004; Şimşek, 2007).

CFA results indicate χ^2=652.27 and df = 523. Therefore, χ^2 / df = 1.24. Chi-square / Degree of Freedom must have a value < 2, which is an indicator of a fairly good relationship or < 5, which is accepted as indicator of good relationship (Harrington, 2008). 1.24 value attained in this study indicates that four-factor construct of the latent variable fits with the 34 observed variable in an acceptable range. On the other hand, Goodness of Fit Index (GFI) and Adjusted Goodness of Fit Index (AGFI) values range from 0.00 to 1.00 and 0.00 stands for no concordance while 1.00 is seen as a sign of perfect concordance. Those GFI and AGFI values at or above 0.90 and 0.85 respectively shows a good concordance (Harrington, 2008). As seen in the table, the model produced results within or close to acceptable ranges (GFI = 0.79; AGFI = 0, 82). Harrington (2008) sets the acceptable ranges for Root Mean Square Error of Approximation (RMSEA) value between. 00 and .08 as an indicator of good concordance while .06 as a cut point. In the present study, RMSEA = 0,082 was found to be close to the edge of acceptable range. On the other hand, Root Mean Square Residual (RMR) and Standardized Root Mean Square Residual (SRMR) values below .05 indicate a good concordance while values below .08 signal acceptable concordance. The related values were shown to be 0.013 indicating a good concordance with the explored model. Finally, Normed Fit Index (NFI), Comparative Fit Index (CFI) and Incremental Fit Index (IFI) must be 0.90 or greater than it. As seen in Table 3.47, NFI, CFI and IFI indices were found to be very close to the acceptable ranges (NFI = 0,85; CFI = 0,87; IFI = 0,88). As a result of the evaluation of the results shown in table 3.47 based on the acceptable ranges for the fit indices specified in the literature (Harrington, 2008; Schumacker, Lomax, 2004; Şimşek, 2007), it can be understood that fit indices attained in the current study are within or very close to the acceptable ranges. As a composite evaluation of fit indices, standardized solution values and t-values obtained from path diagrams, the researcher confirmed EFL self-efficacy Belief Scale explored to include 34 items gathered around four factors as a valid assessor of EFL self-efficacy beliefs among university prep students.

The last and fourth fold of the analysis of the data gathered from the administration of EFL Self-Efficacy Beliefs in the pilot study was devoted to the reliability analysis. In this respect, cronbach's alpha (α) value using *SPSS version 21 for Windows* was computed to test the reliability of the scale since cronbach's alpha, instead of KR_{20}, is the most common measure of internal consistency especially with multiple likert questions in a survey/questionnaire that form a scale and when the aim of the researchers is to determine if the scale is reliable (Bland, Altman, 1997; Cronbach, 1951).

Table 3.48: Reliability Statistics for EFL Self-Efficacy Belief Scale

Number of Items	Cronbach'saplha (α)	$\overline{x}$	S
34	0.94	94	19.65

As a generally accepted rule for describing internal consistency of a scale with likert items using cronbach's alpha value, Cortina (1993) sets the values in the range greater than 0.7 indicate acceptable reliability of measurements for scales consisted of likert items while that of α greater than 0.90 stands for excellent reliability coefficiency. Based on the acceptable ranges of cronbach's alpha for scale with likert items, the confirmed EFL Self-Efficacy Belief Scale was found to be a reliable assessor of EFL self-efficacy beliefs among university prep students (α = 0.94; $\overline{x}$ = 94; s = 19,65).

To sum up, based on the results of CFA and reliability analyses, the adopted scale of self-efficacy including 34 items in four factors was found in the pilot study as a valid and reliable assessor of university prep EFL learners' self-reported efficacy beliefs in four main skills of English (RMSEA=0.080; SRMR=0.013; χ2 / df = 1,24; α = 0.94).

3.7.3. Pilot of Attitudes toward English Scale

In the current study, it was aimed to explore the impact of the flipped classroom treatment on the participant EFL students' attitudes toward English. Therefore, *Attitudes Toward English Scale* developed by Tunç (2003) based on 'Attitudes toward Mathematics and Physics Scale' developed by Aiken (1979) and then

validated by Tuncer, Berkant, Doğan (2015) was administered as pre- and post-tests to gain insights about the impact of the flipped classroom on the learners' attitudes toward EFL (see Appendix 9).The original scale consisted of 24 items scored on 5-point likert type and gathered around four factors: Love for Math, Fear of Math, Motivation for Math and Significance of Math. Under every factor, there were six items, three of which were negative while the remaining three were positive. Tunç (2003) adapted the scale into Turkish in a way to determine university prep EFL learners' attitudes toward English. By being loyal to its original form, Tunç (2003) modified the original four factored scale to match up with EFL (α= .77). Yet, the scale was not validated by Tunç. Elaborate validity and reliability procedures of the scale were conducted by Tuncer, Berkant, Doğan (2015) on the data collected from 271 university prep EFL learners. The results of EFA and CFA explored and confirmed that the scale had four factors with 19 items explaining the 54.180 % of the total variance indicating the instrument as a valid and reliable assessor of university prep EFL learners' attitudes toward English (RMSEA=0.059; SRMR=0.061; α= 88.).

In the pilot study, this scale was administered to test its statistical relevance to the sample of the main study. This started with taking the permission to use the scale for the current study and run CFA by using *LISREL 8.54 for Windows* based on a full structural equation model (SEM) to test and confirm the four-dimensioned model of the adopted scale in the pilot study.

Within this perspective, the researcher took the same steps as those in the pilot of EFL Self-Efficacy Belief Scale. In four folds, the data was exposed to Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) & Bartlett's Test of Sphericity, The Kolmogorov-Smirnov test to come up with the most suitable estimation method, CFA and internal consistency estimate of reliability through cronbach's alpha (α) value through *SPSS 21 for Windows*.

To measure the adequacy of the sampling for factor analysis in the pilot study, *KMO & Bartlett's Test of Sphericity* was computed through *SPSS 21 for Windows*. As a result, the measure of the sampling adequacy was found to be higher than minimum acceptable value set by Kaiser (1960) as greater than 0.5. This KMO and Bartlett's significance < 0,001 show that the sample size in the pilot study, rejecting null

hypothesis, is well suited to conduct the intended factor analysis (KMO = 0,732; x^2 = 434,477; p < .001).

Following *KMO & Bartlett's Test of Sphericity*, the researcher tested the normality of the distribution by employing *the Kolmogorov-Smirnov Test* through *SPSS 21 for Windows* to decide on the best estimation method in CFA. The result of this test of normality was presented in the following table:

Table 3.49: Kolmogorov-Smirnov Test for the Normality of the Scores of Attitudes toward English Scale

		Attitudes toward English
N		38
Parameters	$\overline{x}$	2,6247
	S	,59005
K-Smirnov Z		,670
P		,761

To come up with a decision whether the data distributes normally or not, *Kolmogorov Smirnov Test* was run on the data gathered from the administration of Attitudes toward English Scale in the pilot study. As a result, as seen in the table, the data for the variable was found not to differ significantly from the normal distribution (z =,670; p >.05). In this respect, Maximum Likelihood Method was chosen as the estimation method in CFA since it is reported to be the best and most frequently employed estimation method in CFA when the data is made of continuous variables that distribute normally (Şimşek, 2007).

In the third fold, the statistical relevance of the adopted scale was tested through CFA based on Maximum Likelihood Method using *LISREL 8.54 for Windows*. In this respect, the researcher attempted to confirm how much of the latent variable, attitudes toward English gathered around four sub-factors (attitudes toward learning, individual meaning, motivation and anxiety, importance of English) were explained by 19 observed variables degraded from 24 items of the original scale. Latent variables were shown in ellipses while 19 observed variables were shown in 19

rectangles. The factor loadings of observed variables from Item1 to Item19 on latent variables were indicated in the following two figures.

Chi-Square=173.45, df=148, P-value= 0,0489 RMSEA=0.068
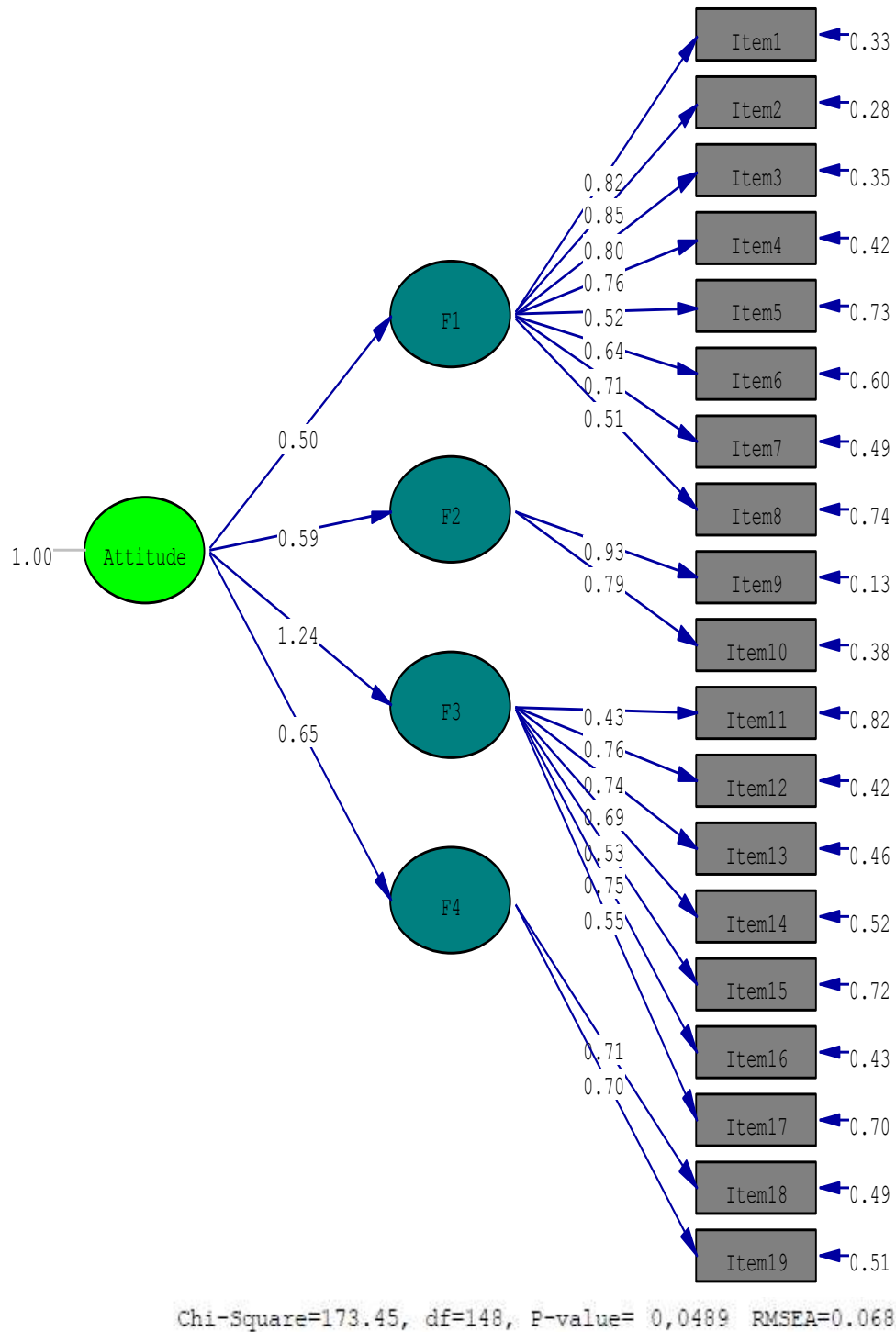
Figure 3.13: Model and Standardized Solutions for Attitudes toward English Scale

In the Figure 3.13, the standardized solution results indicate that the observed 19 variables represent the latent variable (attitudes toward English) in four sub-scales ranging from 0.43 to 0.93. These results confirm the four-sub-scaled model validated by Tuncer, Berkant, Doğan (2015) in 19 items with extraction values ranging 0.52 to 0.80 explaining the 54.180 % of the total variance.

As seen in the figure, the confirmed model produces a perfect match with the model. In other words, the first 8 observed variables were found to explain F1 (attitudes toward learning) while 2 observed variables were confirmed to explain F2 (individual meaning loaded in English). On the other hand, 7 variables from 11 to 17 were found to explain F3 (motivation and anxiety) while the last 2 variables were found to explain F4 (importance of English).

Harrington (2008) and Şimşek (2007) suggest that factor loadings of an observed variable in a confirmed model should be no less than 0.30 and error covariance values have to be low to better explain a latent variable. In this respect, the factor loadings of all the items in the piloted scale were found to be acceptable ranging from 0,43 to 0,93. On the other hand, error covariance values, which stand for the variance in each measurement that do not covary with the latent variable, were also found to well explain the latent variable in acceptable ranges of error. When the error covariance values and factor loadings values are analyzed together, it was decided to keep all the items in the model as confirmed by Tuncer, Berkant, Doğan (2015).
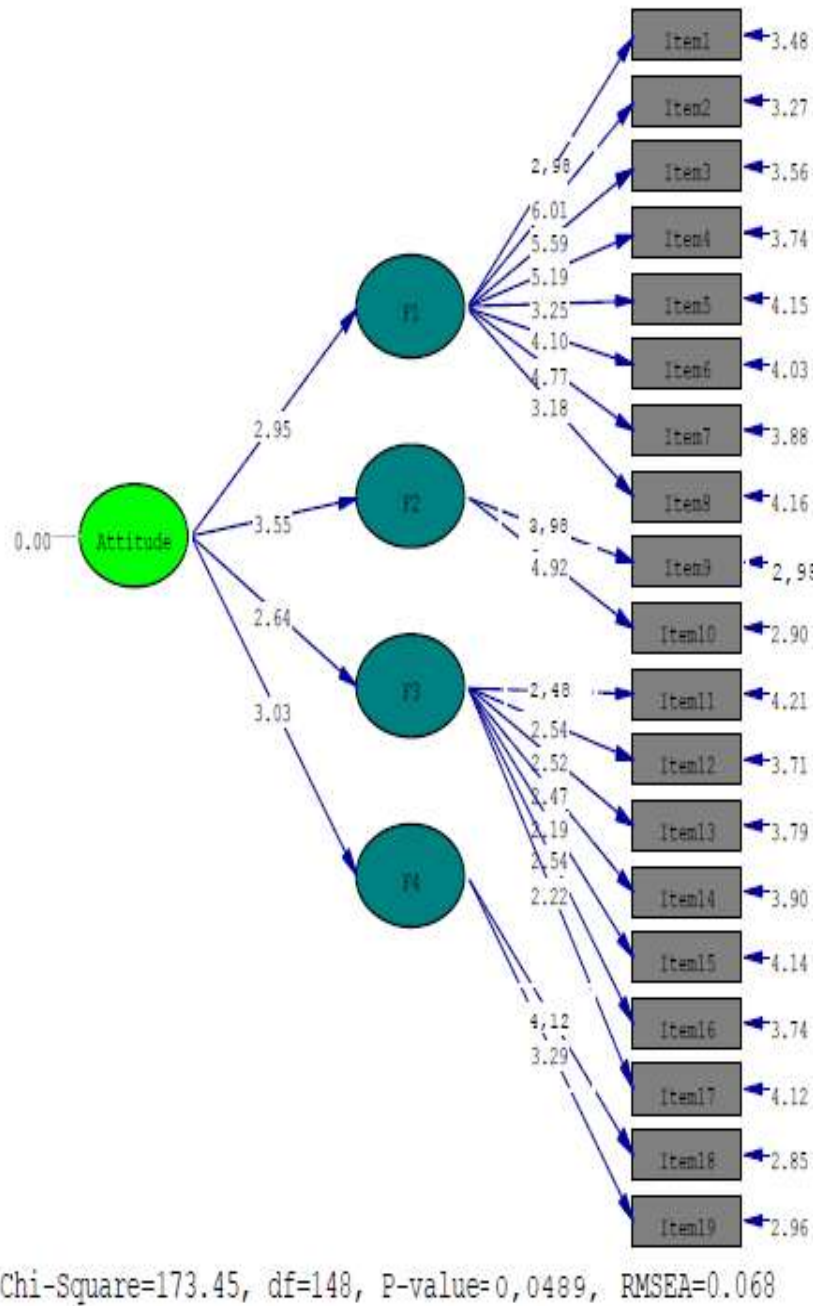
Figure 3.14: Not Standardized T-Values for Attitudes toward English Scale

T-values in Figure 3.14 range from 2.22 to 6.01. Statistically, when the significance .05 is accepted as a criterion in social sciences, t-values must be ≥1.96 while it must be ≥ 2.56 at 0,001 significance level (Şimşek, 2007; Çapık, 2014). Therefore, the variables having t-value less than 1.96 are advised to be omitted from the scale. However, no variable, as seen in the figure, has a t-value below 1.96. Moreover, 12 of the observed variables have t-values greater than 2.56 (p < 0.001) while 7 of them have t-values in the range from 2.22 to 2.54 (p < 0.05). As a result of these not

standardized t-values, the adopted scale is confirmed to be a valid predictor of university prep students' EFL self-efficacy beliefs.

In addition to standardized solutions and t-values obtained from path diagrams as shown in the figures above, the researcher also checked fit indices to confirm the scale as a valid assessor of EFL self-efficacy beliefs in four dimensions. Therefore, the following table presents the necessary fit indices related to the model.

Table 3.50: Fit Indices for Attitudes toward English Scale

Fit Indices	Values
x2	225.18
Df	148
Chi-square (x2) / df	1.52
GFI	0.87
AGFI	0.88
CFI	0.90
RMSEA	0.068
SRMR	0,012
RMR	0,012
NFI	0.76
IFI	0.90

To make a stronger conclusive remark on the scale, the values in Table 3.49 were evaluated following the acceptable ranges specified in the literature (Harrington, 2008; Schumacker, Lomax, 2004; Şimşek, 2007).

CFA results indicate χ^2 / df = 1.52 standing for a value < 2, which is an indicator of a fairly good relationship. On the other hand, Goodness of Fit Index (GFI) and Adjusted Goodness of Fit Index (AGFI) values range from 0.00 to 1.00 and 0.00 stands for no concordance while 1.00 is seen as a sign of perfect concordance. GFI and AGFI values in the table indicate good concordance with the model (GFI =0.87;

185

AGFI = 0.88). In the present study, RMSEA = 0,068 was found to be in the acceptable range. On the other hand, Root Mean Square Residual (RMR) and Standardized Root Mean Square Residual (SRMR) values were shown to be 0.012 indicating a good concordance with the explored model. Finally, Normed Fit Index (NFI), Comparative Fit Index (CFI) and Incremental Fit Index (IFI) must be 0.90 or greater than it. As seen in the table, NFI, CFI and IFI indices were found to be very close to or in the acceptable ranges (NFI = 0.76; CFI = 0.90; IFI = 0,90). As a result of the evaluation of the results shown in Table 3.50 based on the acceptable ranges for the fit indices specified in the literature (Harrington, 2008; Schumacker, Lomax, 2004; Şimşek, 2007), it can be understood that fit indices attained in the current study are within or very close to the acceptable ranges. As a composite evaluation of fit indices, standardized solution values and t-values obtained from path diagrams, the researcher confirmed Attitudes toward English Scale was explored to include 19 items gathered around four factors as a valid assessor of EFL attitudes toward English among university prep students.

The last and fourth fold of the analysis of the data gathered from the administration of EFL Self-Efficacy Beliefs in the pilot study was devoted to the reliability analysis. In this respect, cronbach's alpha (α) value using *SPSS version 21 for Windows* was computed to test the reliability of the scale since cronbach's alpha, instead of $KR_{20,}$ is the most common measure of internal consistency especially with multiple likert questions in a survey/questionnaire that form a scale and when the aim of the researchers is to determine if the scale is reliable (Bland, Altman,1997; Cronbach, 1951).

In the last and fourth fold of the analysis, cronbach's alpha (α) value using *SPSS version 21 for Windows* was computed to test the reliability of the scale.

Table 3.51: Reliability Statistics for Attitudes toward English Scale

Number of Items	Cronbach'saplha (α)	$\overline{x}$	S
19	0.91	49.8	11.21

Based on the acceptable ranges of cronbach's alpha for scale with likert items, the confirmed Attitudes toward English Scale was found to be a reliable assessor of university prep students' attitudes toward English ($\alpha = 0.91; \overline{x} = 49.8 ; s = 11.21$).

3.7.4. Pilot of Semi-Structured Interviews

Following the aforementioned procedures in the pilot study, semi-structured interviews were held. To seek for participants' perceptions about their flipped classroom experiences and its impact on their attitudes toward and self-efficacy beliefs in EFL, the researcher came up with a semi-structured form consisted of ten questions (see appendix 11). Preceding the pilot study, the procedures to justify and validate the form were realized with the participation of seven experts from different subject areas in different phases. These procedures were reflected in part 3.3.4 in detail. As a result, the form was validated with the mean score of 4,8 out of 5 indicating the semi structured form as a completely valid assessor of EFL students' perceptions of flipped classroom model in EFL. Following the *multiple variation sampling* method, the interviews were held with 5 voluntary students in the flipped EFL classroom from three different achievement groups as low, medium and high. The pilot of interviews was held not to produce qualitative data to be analyzed to support quantitative data in the pilot study but to find out unexpected problems that may hinder the research procedures (Van Teijlingen, Hundley). With an awareness of exploring any problems that may come out during these semi-controlled protocols to shape the main procedures in the major study, the interviews were held individually at university at the consensual time. As a result of this procedure, the interviews were found to last 25 to 40 minutes. During the interview, it was seen that the interview questions were clearly understood by the interviewee and the place where the interviews were realized and the anticipated time allocated for covering them were found to be adequate.

To conclude, the pilot study provided invaluable opportunities to help the researcher draw the framework of the current study with an awareness of the probable problems that may come out in the study and destroy its logical flow. Ensuring both quality and efficiency of the study, the pilot study justified and validated the instruments to be used in the main study as a valid and reliable means of collecting data in response to the purpose of the study.

4. FINDINGS

This embedded study grounded in a mixed method research design was carried out in order to explore the impact of flipped classroom teaching method on the university prep EFL learners' academic performance and its predictors such as their attitudes toward and self-efficacy beliefs in EFL. With this purpose in mind, the researcher drew upon both qualitative and quantitative data to better understand the issue. In this respect, the data for the study was collected from sources: pre- and post-administration of 3 surveys to the sample of the research study before and after the flipped intervention and follow-up semi-controlled interviews with the experiment group students about their experiences as a facilitator of the preceding quantitative outcomes. As a result, this required the researcher to employ mixed method data analysis techniques based on blending of qualitative and quantitative methods in a single study.

The current mixed method study was conducted in two phases where quasi-experimental pre- and post-test design was followed by the phenomenological design. The first experimental design referred to quantitative analysis including *Independent Sample T* and *Paired Sample T Tests* to explore the impact of the flipped intervention on the participants' EFL performance and its predictors. On the other hand, the phenomenological design appealed to content analysis procedures to support and enrich the results of the study by eliminating any potential bias that the quantitative data could bring into the study.

To conclude, this chapter reports the results of the analysis of the data collected in the main study in folds. While the first part covers the attempts to determine the type of the data to employ best analysis techniques, the second part stands for the revision of the analysis reported in the part about "participants" to equate control and experimental groups based on their pre-tests scores of 3 data collection instruments of the study. These analyses, not only to enrich and vary the techniques of data analysis but also avoid the risk of type 1 error, were revised applying MANOVA instead of independent sample t tests. The third fold included the comparison of pre-

and post-test scores of the participants within their groups while the following fold covered the analyses that indicate the difference in post-test scores and retention of achievement test scores between groups. Following these analyses that answer the main questions of the study, the last fold, on the other hand, was organized to display the results of content analysis to back up quantitative results.

4.1. Statistics for Identifying the Type of the Data

This part of the study was dedicated to the attempts to adopt the best ground where the data analysis methods and techniques were situated. There are two types of test data as parametric and non-parametric that require different types of analysis. Therefore, identification of the type of the data stands as a prerequisite to adopt the best techniques to come up with valid results. This will naturally donate the researcher with the ability to draw reliable and valid conclusions about the sample. With this goal in mind, the researchers checked the total data in terms of the number of participants, measures of central tendency, normal distribution curve and the results of the normality test. In addition to testing the normality of distribution regarding participants' total scores, the same procedures were followed to compute if the data distributes normally within the groups to strengthen the reliability of the analyses.

Within this respect, the following table reports the results of checking the data related to 3 variables of the study in terms of the number of participants, measures of central tendency.

Table 4.1: Measures of Central Tendency for Total Scores

	Pre-testing			Post-testing		
	Efficacy	**Attitude**	**EFL Achievement**	**Efficacy**	**Attitude**	**EFL Achievement**
N	41	41	41	41	41	41
Mean	2,725	3,248	48,54	2,984	3,1502	67,21
Median	2,696	3,263	51	3,006	3,2632	68
Mode	2,3	3,2	56	1,8	3,11	78
Skewness	,156	-,407	-,378	-,163	-1,212	-1,022
Kurtosis	-,765	,666	-,542	-,016	2,128	1,389

As shown in the table, the number of participants with regard to each of the variables of the study are above 30 (n= 41). However, the number of the participants within the groups are below 30 (n= 21; n= 20). Based on *Central Limit Theorem (CLT)*, Kul (2014) and Marasinghe et al (1996) state that the samples whose sizes are n > 30 are assumed to be close to distribute normally. Therefore, the total number of the participants indicates normal distribution, while their numbers in control and experimental groups do not. Yet, CLT is not a rule of thumb to determine the normality of the distribution on its own. If the assumptions of normality are severely violated, at that time nonparametric tests must be used (Turner & Thayer, 2001). In this respect, ultimate decision was made after evaluating all the assumptions regarding the distribution of data within control and experimental groups.

As another basis of evaluation, the measures of central tendency related to each of the variables of the study were taken into account. Balcı (2005) indicates that in a normal distribution, the measures of the set of data get approximate values to the mean. In this respect, mean, median and mode with regard to total scores of each of the variables in pre and post-testing are shown to be very close to each other. While the mean of pre-test scores of Self-efficacy Scale was reported to be 2,725, median and mode of that was found to be respectively 2,699 and 2,3. The same close range was observed in the pre-test scores of Attitude Scale and EFL Achievement test respectively (Mean= 3,248, Median= 3,263, Mode= 3,2; Mean= 48,54, Median= 51,

Mode= 56). The close range of measures in the post-testing also indicates the normality of the distribution with regard to three variables of the study.

On the other hand, as another indicator for the normality of the distribution of the data, the values of *skewness and kurtosis* related to total scores of 3 variables of the study in pre and post-testing were investigated. There seems to be no rule of thumb to follow to interpret the results of skewness and kurtosis for normal distribution. However, it is advised to follow indices for acceptable limits of ± 3 (Gravetter, Wallnau, 2014). As seen in the table, values of *skewness* with regard to efficacy, attitude and achievement variables in pre (,156; -,407; -,378) and post-testing (-,163; -1,122; -1,022) were found to be within acceptable range respectively. On the other hand, the values of *kurtosis* were also found to stand within acceptable limits in both pre (-,765; ,666; -,542) and post-testing (-,016; 2,128; 1,389) respectively.

Additionally, further evaluation the values of central tendency including the values of *skewness and kurtois* were held related to the data within control and experimental groups. The findings were shown in the following two tables.

Table 4.2: Measures of Central Tendency for the Control Group

	Pre-testing			Post-testing		
	Efficacy	**Attitude**	**EFL Achievement**	**Efficacy**	**Attitude**	**EFL Achievement**
N	20	20	20	20	20	20
Mean	2,948	3,253	50,65	2,788	2,9000	62,15
Median	2,996	3,289	55,00	2,886	3,2105	63,50
Mode	2,1	3,4	31	1,8	2,21	61
Skewness	,109	,095	-,356	-,780	-1,004	-,490
Kurtosis	-,648	-,767	-,800	-,272	-,143	,057

As seen in the table, regarding the assumptions of measures of central tendency, the data within the control group was found to distribute normally by having median and mode values close to the mean scores of three variables with skewness and kurtois indices within acceptable ranges.

191

Table 4.3: Measures of Central Tendency for the Experimental Group

	Pre-testing			Post-testing		
	Efficacy	Attitude	EFL Achievement	Efficacy	Attitude	EFL Achievement
N	21	21	21	21	21	21
Mean	3,171	3,243	46,52	2,596	3,4313	72,05
Median	3,279	3,263	46,00	2,381	3,4211	72,00
Mode	1,8	3,1	46	1,5	3,11	78
Skewness	-,234	-1,418	-,420	,216	,608	-,056
Kurtosis	-,146	2,040	-,245	-,775	-,297	-1,337

Table 4.3 indicates that the data within the experimental group have close measures of central tendency in addition to skewness and kurtosis values within acceptable ranges.

In addition to measures of central tendency, the following part stood for evaluation of the data with regard to 3 different variables checked through normal distribution curve. Normal distribution is sometimes referred as *normal distribution curve*, *Gauss distribution* or *Gauss bell curve* in the related literature (Lyon, 2014). Based on *CLT,* it is theoretically set that "many variables that we see in nature appear to have a probability density function that approximates a bell–shaped curve" (Lyon, 2014, 622). As a result, all the data with regard to three variables of the study were found to distribute normally. The following normal distribution curves are presented to show distribution of total pre and post test scores with regard to 3 variables of the study.
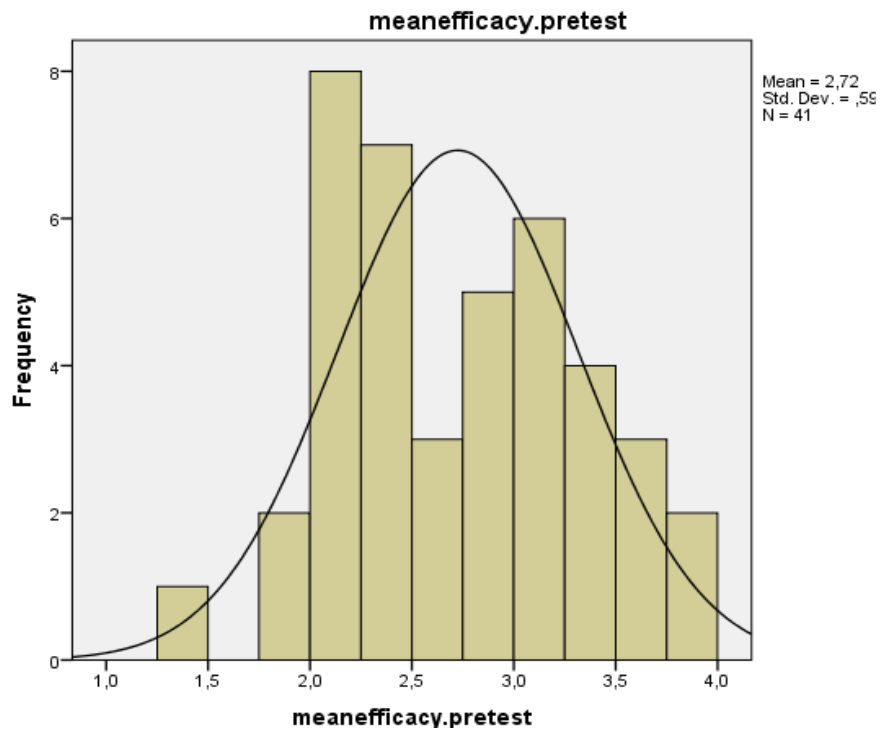
Figure 4.1: Normal Distribution Curve for Pre-test Scores of Self-Efficacy
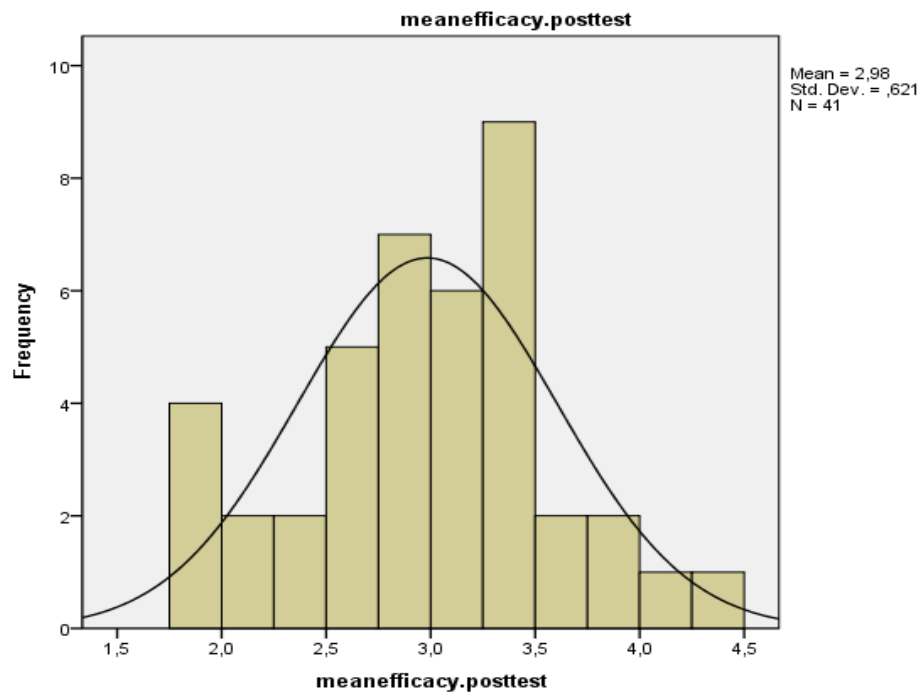
Figure 4.2: Normal Distribution Curve for Post-test Scores of Self-Efficacy

As seen in the figures, the self-efficacy scores of the participant EFL students get closer to normal distribution with no bias left or right centering around its mean. In

addition to the curves for the participants' pre and post-test scores of self-efficacy beliefs in EFL scale, following figures are set to indicate how those of attitudes toward EFL distribute.
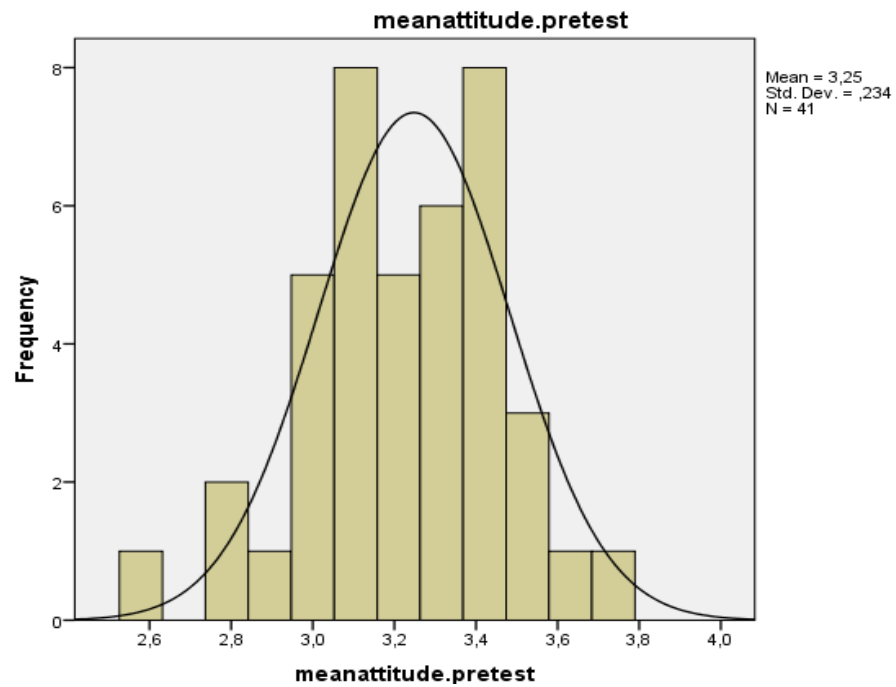
Figure 4.3: Normal Distribution Curve for Pre-test Scores of Attitudes
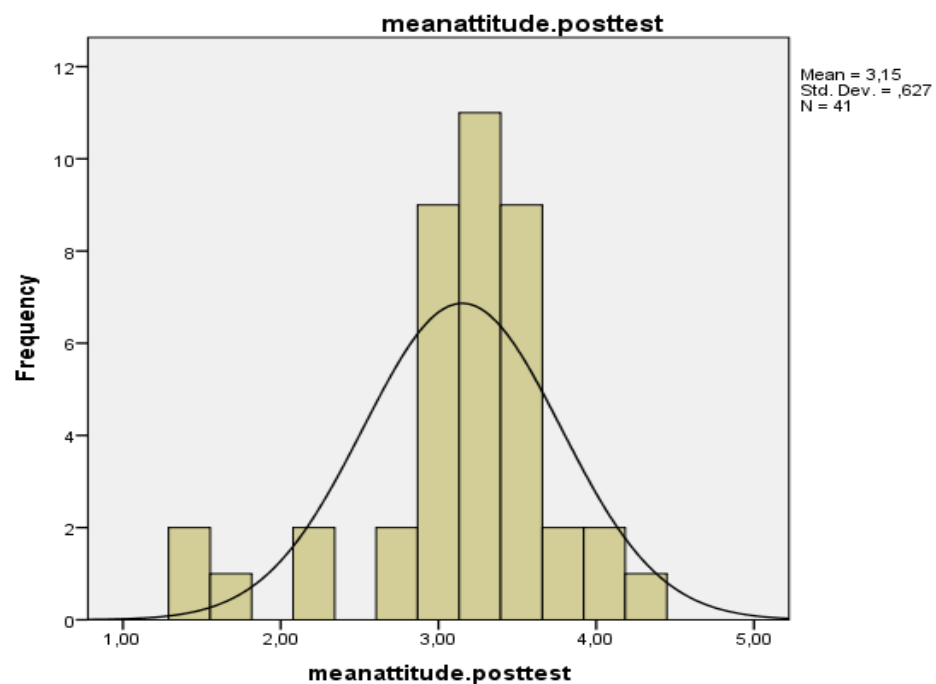
Figure 4.4: Normal Distribution Curve for Post-test Scores of Attitudes

194

As seen in the Figures 4.3 and 4.4 mean of the scores the participant student obtained from *Attitudes toward EFL Scale* got changed into single-mode normal distribution from pre to post intervention of the scale. Last two figures below show normal distribution curves for EFL Achievement pre and post-test scores.
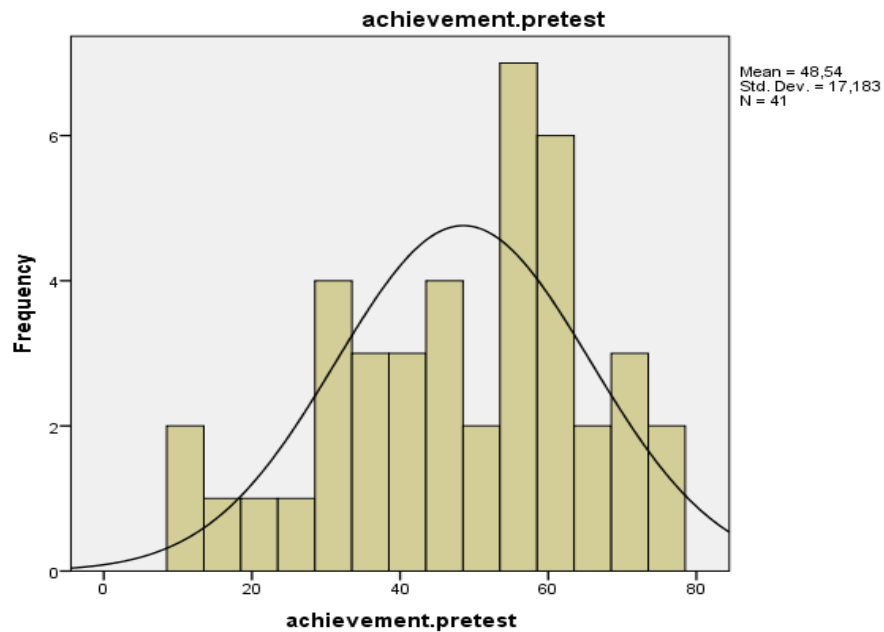
Figure 4.5: Normal Distribution Curve for EFL Achievement Pre-test Scores
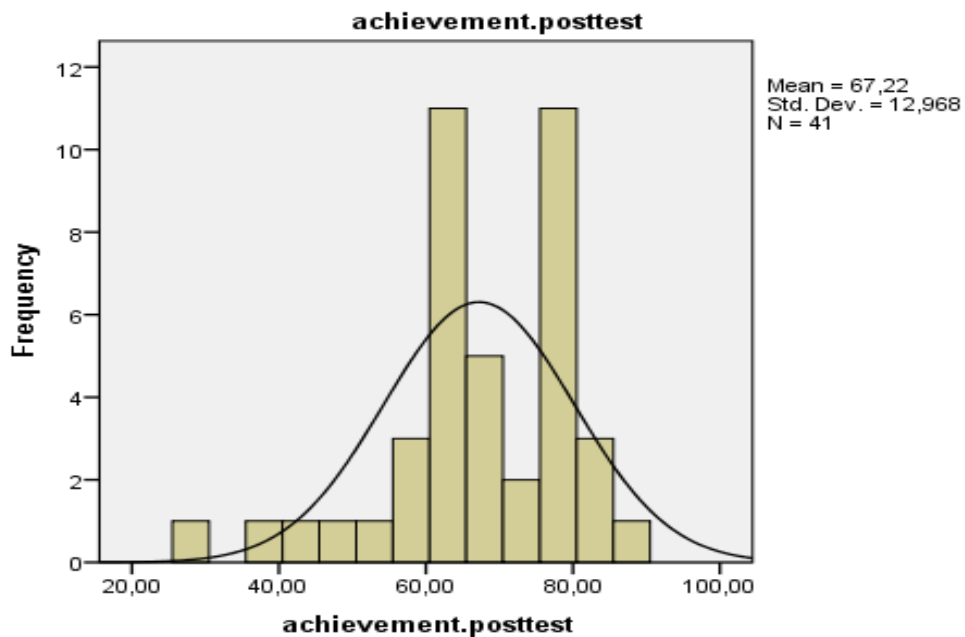
Figure 4.6: Normal Distribution Curve for EFL Achievement Post-test Scores

Figures 4.5 and 4.6 indicate that EFL Achievement pre-test scores distribute normally but a bit positively skewed. However, this does not make it significantly different than normal distribution as also indicated by the values of central tendency, skewness and kurtosis. On the other hand, EFL Achievement post-test scores were shown to distribute more normally with decreasing positive skewness on the right.

This step was further expanded by checking the normal distribution curves for the data within the groups. Not to cover much of the space, they were not documented in the body of the study. Some of the curves within the groups were found to be a bit positively skewed. Yet, relying on the values of central tendency, skewness and kurtosis that range close to each other within acceptable limits, it was explored that normal distribution curves within control and experimental groups are not significantly different than normal distribution.

In the last step of deciding whether the data distributes normally or not to adopt the best analysis methods to come up with valid results, six separate *Kolmogorov Smirnov tests* were run for the data obtained as a result of pre- and post-intervention of three data collection instruments. *Kolmogorov Smirnov test* was adopted since it is the statistical test employed to determine if a sample comes from a population with a normal distribution by measuring the distance between the empirical distribution function of the sample and the cumulative distribution function of the specific distribution (Lilliefors, 1967).

Table 4.4: Kolmogorov Smirnov Tests for the Normality of the Distributions for Total Scores

		Pre-testing			Post-testing		
		Efficacy	Attitude	EFL Achievement	Efficacy	Attitude	EFL Achievement
N		41	41	41	41	41	41
Parameters	$\overline{\text{x}}$	2,725	3,248	48,54	2,98	3,1502	67,22
	S	,5904	,2343	17,183	,6212	,62691	12,96
K-Smirnov Z		,781	,559	,721	,519	1,296	,772
p		,576	,913	,677	,951	,069	,590

As a result, as seen in the table, the data for the variables obtained from pre ($z=$,781; ,559; ,721) and post testing ($z=$,519; 1,296; ,772) were found not to differ significantly from the normal distribution (p >.05)

To further validate the analyses to be run in the study, the researcher computed *Kolmogorov Smirnov tests* to set the normal distribution of the data within the control and experimental groups. Hence, separate tests were run to refer to data regarding each of the variables in pre- and post-testing.

Table 4.5: Kolmogorov Smirnov Tests for the Normality of the Distributions for the Control Group

		Pre-testing			Post-testing		
		Efficacy	Attitude	EFL Achievement	Efficacy	Attitude	EFL Achievement
N		20	20	20	20	20	20
Parameters	$\overline{\text{x}}$	2,948	3,253	50,65	2,788	2,9000	62,15
	S	,5028	,2632	17,027	,5289	,73821	15,513
K-Smirnov Z		,544	,621	,572	,595	,844	,539
p		,929	,835	,899	,871	,475	,934

Table 4.5 shows that the scores obtained from three data collection instruments in pre testing ($z=$,544; ,621; ,572) and post testing ($z=$,595; ,844; ,539) were found not to differ significantly from the normal distribution (p >.05)

Table 4.6: Kolmogorov Smirnov Tests for the Normality of the Distributions for the Experimental Group

		Pre-testing			Post-testing		
		Efficacy	Attitude	EFL Achievement	Efficacy	Attitude	EFL Achievement
N		21	21	21	21	21	21
Parameters	$\overline{x}$	2,596	3,243	46,52	3,171	3,4313	72,05
	S	,6114	,2095	17,503	,6563	,31911	7,586
K-Smirnov Z		,738	,733	,430	,502	,630	,802
P		,648	,635	,993	,963	,822	,541

As seen in the table, the data obtained from pre (*z*=,738; ,733; ,430) and post testing (*z*=,502; ,630; ,802) within the experimental group were found not to differ significantly from the normal distribution (p >.05)

As a result of all these analyses run in three folds, the data obtained in pre and post intervention of the instruments related to three different variables of the study were checked regarding total scores and the scores within the groups. These elaborate analyses found that the data distributed normally both totally and within the groups. The only problem was about the number of participants below 30 in control and experimental groups. However, relying on the fact that all other assumptions were met and the suggestion by Turner and Thayer (2001) that if assumptions of normality are not severely violated, parametric tests be used, it was decided to employ parametric tests in order to answer the research questions of the study. In this respect, *Independent Sample T* and *Paired Sample T Tests* were run *to* detail the impact of the flipped intervention on the participants' EFL performance and its predictors between and within groups.

There seems no thumb of rule that determine the minimum sample size for t tests to be valid. "The t-test assumes that samples are randomly drawn from normally distributed populations with unknown population means" (Park, 2009, 1). The problem that may occur with low sample size is the strongest characteristic of the test. In other words, producing improved results for small sized samples is t test's claim to realize. To support this claim, the very first application of the t-test was

demonstrated in sample sizes of size four in 1908 (Student, 1908). In specific terms, number of participants in control and experimental groups below 30 (N= 20, 21) is not a fact that hinders the application of t tests in the present study since there seems no rule of thumb to follow regarding the number of participants in the literature.. In this respect, as an exmple, "the Central Limit Theorem says, however, that the distributions of y_1 and y_2 are approximately normal when N is large. When $n_1 + n_2 \geq$ 30, in practice, you do not need to worry too much about the normality assumption" (Park, 2009, 2). If the sample size is small and the null hypothesis of normality is rejected, *Kolmogorov-Smirnov test* is suggested by Park (2009) to employ to test normality assumption. Within all these circumstances, t test is explored to be statistically applicable to the data set in this present study.

Dwelling on all the analyses that indicate the statistical normal distribution of data, the researcher took the next step to check if the data obtained from pre- and post-testing of each scale meet the assumptions for t test. Within this perspective, the review of related literature indicates five main assumptions must be met independent samples and paired sample t tests to produce valid results (Büyüköztürk, 2007a; Creswell, 2012; Karasar, 2007):

1. Dependent variable must be measured on a continuous scale.

2. Independent variable must include two independent sub-categories.

3. Sub-groups must consist of different participants.

4. Dependent variable must approximately distribute normally for each group of independent variable.

5. Variances of dependent variable for each group of independent variable must be homogenous.

The detailed analyses of the data within control and experimental groups for pre- and post-test scores obtained regarding each dependent variable in the previous section found out that the data met the first four assumptions for the administration of t-tests. After exploring that the dependent variables were measured on three different continuous scales at ratio or interval level with the participation of two independent group of students that produced normally distributed data, the researcher computed Levene's tests for the homogeneity of variances for each group of variable. Levene's test is used to verify that *k* samples have equal variances (Gastwirth, Gel, Miao,

2009). The results of separate Levene's Tests were presented below to assess the homogeneity of variances across groups regarding pre- and post-test scores for each dependent variable below:

Table 4.7: Homogeneity of Variances for the Scores of EFL Achievement Test

Factor	Group	Pre-Testing		Post-testing	
		Levene's Test		Levene's Test	
		F	*p*	*F*	*P*
EFL Listening	Control	,002	,963	1,838	,183
	Exper.				
EFL Grammar	Control	,054	,818	2,321	,136
	Exper.				
EFL Reading	Control	,715	,403	1,030	,317
	Exper.				
EFL Vocabulary	Control	,015	,902	,134	,716
	Exper.				
EFL Writing	Control	1,699	,200	,196	,661
	Exper.				
Total EFL Test Scores	Control	,006	,937	3,538	,067
	Exper.				

The elaborate discussion for parametric tests already done in previous parts formed the first fold of the assumptions met. In the second fold, the researcher checked the data for the homogeneity of variances through separate *Levene's Test.* In this respect, the table expands the inquiry about the assumptions for T-tests regarding the data obtained from each sub-scale of EFL Achievement Test. As a result, as seen in the table, the variances of pre- (F = ,002; ,054; ,715; ,015; 1,699; ,006) and post-test scores (F = ,183; ,136; ,317; ,716; ,196; 3,538) obtained from achievement test across groups were found to be homogenous (p >.05).

Table 4.8: Homogeneity of Variances for the Scores of Attitudes toward EFL Scale

Factor	Group	Pre-Testing		Post-testing	
		Levene's Test		Levene's Test	
		F	*p*	*F*	*p*
Attitudes Toward EFL	**Control**	2,753	,105	3,362	,074
	Exper.				

In addition to four assumptions covered above, the researcher checked the data obtained from the pre- and post-test administration of Attitudes toward English Scale for the homogeneity of variances. In this respect, *Levene's Tests* indicated the homogeneity of variances across groups (*F* = 2,753; 3,362; p >.05).

Table 4.9: Homogeneity of Variances for the Scores of Self-Efficacy Belief in EFL Scale

Factor	Group	Pre-Testing		Post-testing	
		Levene's Test		Levene's Test	
		F	*p*	*F*	*p*
Self-Efficacy in EFL Reading	**Control**	,550	,463	,326	,571
	Exper.				
Self-Efficacy in EFL Writing	**Control**	,558	,459	,037	,849
	Exper.				
Self-Efficacy in EFL Listening	**Control**	,078	,781	4,348	,053
	Exper.				
Self-Efficacy in EFL Speaking	**Control**	,121	,731	,808	,374
	Exper.				
Total Self-Efficacy in EFL Scores	**Control**	,374	,544	,795	,378
	Exper.				

As seen in the table, expanding the normality-specific analyses, the researcher computed Levene's Test for the homogeneity of variances for pre- and post-test scores obtained from Self-Efficacy Belief EFL Scale. As a result, the variances across groups were found to be homogenous for pre-test (F = ,550; ,558; ,078; ,121; ,374; p >.05) and post-test scores (,326; ,037; 4,348; ,808; ,795; p >.05).

Table 4.10: Homogeneity of Variances for the Retention of EFL Achievement Test

Factor	Group	Levene's Test	
		F	*P*
EFL Listening	**Control**	2,654	,111
	Exper.		
EFL Grammar	**Control**	,006	,939
	Exper.		
EFL Reading	**Control**	2,594	,115
	Exper.		
EFL Vocabulary	**Control**	1,763	,192
	Exper.		
EFL Writing	**Control**	1,911	,175
	Exper.		
Total EFL Test Scores	**Control**	3,496	,069
	Exper.		

The last fold of Levene's Tests were run for the variances of Retention of EFL Achievement test between groups. As seen in the table, the variances of the scores were found not to differ significantly (F = 2,654; ,006; 2,594; 1,763; 1,911; 3,496; p >.05).

Dwelling on all these analyses computed to test the assumptions for the normality of distribution and T-tests, the valid and reliable basis was formed to evaluate the differences between and within groups.

4.2. Statistics for Equating the Control and Experimental Groups

This part of the study reflected the results of analyses employed to equate control and experimental groups based on their pre-tests scores of EFL Achievement Test, Attitudes toward EFL and Self-Efficacy Beliefs in EFL scales. Based on the results of analyses that scientifically indicate the normal distribution of data, three separate Independent Samples T-Tests were run to explore if there was any significant difference between two groups of students in terms of the pre-test application results of EFL Achievement Test, Attitudes toward EFL and Self-Efficacy Beliefs in EFL scales.

On this ground, the results of three Independent Samples T-Tests, each standing for a different variable, were shown below:

Table 4.11: Independent Samples T-Tests for Equating the Groups

Factor	Groups	N	$\overline{\text{x}}$	S	$Sh_{\overline{x}}$	t	Df	p	d
EFL Achievement Pre-test	**Control**	20	50,65	17,03	3,81	0,765	39	,449	----
	Exper.	21	46,52	17,50	3,82				
Attitudes Pre-test	**Control**	20	3,25	0,263	,058	0,129	39	,898	----
	Exper.	21	3,24	0,209	,045				
Self-Efficacy Pre-Test	**Control**	20	2,86	0,550	0,12	1,452	39	,154	----
	Exper.	21	2,59	0,611	0,13				

As shown in the table, separate Independent Samples T-Tests were employed to find out the difference between two groups of EFL learners. The results indicated that the students did not differ significantly in terms of the mean pre-tests scores of concerned variables of the study (t = 0765; 0,129; 1,452; p > .05). In other words, it was explored that two groups were assumed to be similar based on their pre-test results. This result was also supported by the fact that they were the students from two different classes at the same proficiency level determined by the placement test held by the university itself. As a result, these two groups of students among whom there was no significant difference in terms of three variables of the study were

assigned as control and experiment groups. Based on these results, the participant students in control and experimental groups were found to be equal regarding their pre-test scores of EFL Achievement Test, Attitudes toward EFL and Self-Efficacy Beliefs in EFL scales. In this way, the researcher ensured the basic ground on which the first phase of the study, quasi-experimental pre- and post-test design.

4.3. Statistics for Pre- and Post-Test Mean Scores of the Control Group

This part of the study covered the statistics that investigate the impact of traditional lecture-based teaching on EFL learners' academic achievement and its sub-sections, their attitudes toward and self-efficacy beliefs in EFL. To accomplish this, the researcher computed separate *Paired Sample T Tests*, also known as *Dependent Sample T Tests*, to explore whether mean scores of the students in the control group related to each of the independent variables (academic achievement, attitude toward and self-efficacy belief in EFL) differ significantly from pre- to post-test.

Within this respect, in order to provide the body of findings that answer the first research question of the study, the researcher presented the results of *Paired Sample T Tests* in three separate tables. While the first table covers the results about the scores of whole achievement test and its sub-sections, the second table investigates those about attitudes toward EFL within the students in the control group. On the other hand, the last of the tables focused on the results of *Paired Sample T tests* about total scores of whole self-efficacy beliefs in EFL scale and its sub-scales.

Table 4.12: Paired Samples T-Tests for EFL Achievement Test

Factor	Group	N	$\overline{x}$	S	$Sh_{\overline{x}}$	t-Test			
						t	*Df*	*p*	*d*
EFL Listening	Pre-test	20	12,05	3,268	,731	-3,106	19	,006	-0,79
	Post-test	20	14,70	3,373	,754				
EFL Grammar	Pre-test	20	9,75	4,756	1,063	-5,160	19	,000	-1,14
	Post-test	20	14,40	3,235	,723				
EFL Reading	Pre-test	20	10,30	4,543	1,016	-3,052	19	,007	-0,61
	Post-test	20	12,75	3,932	,879				
EFL Vocabulary	Pre-test	20	5,90	3,986	,891	-2,557	19	,019	-0,52
	Post-test	20	8,10	4,412	,986				
EFL Writing	Pre-test	20	12,65	4,416	,987	,616	19	,545	---
	Post-test	20	12,20	4,047	,905				
Total EFL Test Scores	Pre-test	20	50,65	17,027	3,807	-4,266	19	,000	-0,70
	Post-test	20	62,15	15,513	3,469				

Based on the results of the multiple analyses that set the data distribute normally, *Paired Samples t tests*, a parametric test, were computed to compare repeated measures of EFL achievement test within the students who were taught following traditional lecture-based instruction model. In other words, Table 4.8 summarizes the findings that indicate the impact of traditional lecturing on the EFL test performance of the students in the control group from pre- to post-test. Within this respect, the results indicate that control group students' post-test scores differed significantly from their pre-test scores in terms of total EFL achievement test (t = -4,266; p <.001) and its four sub-sections as listening (t = -3,106; p <.01), grammar (t = -5,160; p <.001), reading (t= -3,052; p <.01) and vocabulary (t = -2,557; p <.05). Moreover, proved by the mean scores in the table, all the significant differences were in the

favor of post-test scores. However, as seen in the table, the students' scores of writing section did not differ significantly from pre- to post-test (t = 616; p >.05).

In the second fold of these analyses shown in the table, the researcher made an effort to present the effect size of significant difference between EFL achievement pre- and post-test scores. Therefore, Cohen's *d* (degree of impact), which indicates the standardized difference between means for t-tests results regardless of the variables, sample sizes or measurement types (Cohen, 1994), was computed to elaborate on each of these significant differences. In the interpretation of objective effect size of the significant differences, the guidelines provided by Cohen in three ranges as "small = 0.2 < *d* < 0.5," "medium = 0.5 < *d* < 0.8,"," and "large = *d* >.80," (Ruscio, 2008). In this respect, *Cohen's d* values suggested moderate effect size of the significant difference in listening, reading, vocabulary and total EFL test achievement scores within the students in the control group (.*50<d< .*80). On the other hand, following Cohen's (1994) suggestion of ignoring the positive or negative that are caused by positive or negative t value and taking the absolute value, *Cohen's d of – 1,14* indicated large effect of the significant difference in grammar scores of the students instructed traditionally from pre- to post-test (*d>* .80).

Following the analyses that investigate the difference in the means of the repeated measures regarding EFL achievement test and its five sub-sections within the students in the control group, next table reports the findings that compare the difference in the students' mean scores of Attitudes toward EFL scale in the control group from pre- to post-test.

Table 4.13: Paired Samples T-Tests for Attitudes toward EFL

Factor	Group	*N*	$\overline{\text{x}}$	**S**	$Sh_{\bar{x}}$	*t*	*Df*	*p*	*d*
Attitudes Toward EFL	Pre-test	20	3,253	,2632	,0589	2,805	19	,011	0,64
	Post-test	20	2,90	,7382	,1651				

As seen in Table 4.13, the paired sample t test produced results that question the difference in students' mean scores of attitudes toward EFL scale within the control

group from pre- to post-test. Within this perspective, the table reveals that the means of the scores the students who were taught through a traditional lecture-based model got from attitudes toward EFL scale differed negatively and significantly in the favour of pre-test (t = 2,805; p <.05). In other words, better mean of pre-test (3,253) than post-test (2,90) reveals negative impact of traditional teaching approach on the participant students' scores of attitudes toward EFL within the control group.

To detail the impact of the lecture-based instruction on the participant students' scores of Attitude toward EFL scale, *cohen's d* was computed and evaluated on the basis of Cohen's guidelines (1994). As a result of this, Cohen's *d* of 0.64 shows the moderate effect of difference created by the traditional lecture-based model in students' scores of attitudes from pre- to post-test administration (*.50<d< .80*).

In addition to these two tables above, the following table covers the results of five separate Paired Sample T tests to understand the statistical difference in the mean scores of Self-efficacy beliefs in EFL scale within the students in the control group.

Table 4.14: Paired Samples T-Tests for Self-Efficacy Beliefs in EFL

Factor	Group	N	$\overline{\mathrm{x}}$	S	$Sh_{\overline{x}}$	t-Test			
						t	*Df*	*p*	*d*
Self-Efficacy in EFL Reading	Pre-test	20	2,888	,6372	,1425	-,209	19	,837	----
	Post-test	20	2,925	,7951	,1778				
Self-Efficacy in EFL Writing	Pre-test	20	2,800	,5840	,1306	,492	19	,628	----
	Post-test	20	2,705	,6597	,1475				
Self-Efficacy in EFL Listening	Pre-test	20	3,120	,5809	,1299	1,872	19	,077	----
	Post-test	20	2,875	,4423	,0989				
Self-Efficacy in EFL Speaking	Pre-test	20	2,983	,6664	,1490	1,491	19	,152	----
	Post-test	20	2,645	,6312	,1411				
Total Self-Efficacy in EFL Scores	Pre-test	20	2,948	,5028	,1124	1,085	19	,291	----
	Post-test	20	2,788	,5289	,1183				

As seen in Table 4.10, five separate *Paired Samples T tests* were computed to delve into the difference in the mean scores of self-efficacy beliefs in EFL scale within the students in the control group. In other words, this part presents the results that add to the significance of the study since it elaborates on the impact of the instruction on students' beliefs in their own capacities to accomplish certain goals not only in EFL as a whole but also in its four sub-skills as reading, writing, listening and speaking. Within this respect, it was explored that traditional lecture-based instruction model did not exert any significant impact on the means of self-efficacy beliefs in neither EFL as a whole (t = 1,085; p >.05) nor such four sub-skill areas as reading (t= -,209; p >.05), writing (t= ,492; p >.05), listening (t = ,1872; p >.05) and speaking (t = 1,491; p >.05). On the other hand, it is important to note that, though not statistically significant, the means of self-efficacy in EFL writing, listening, speaking and total scores decreased from pre- to post-test administration of the scale. In other words, not in a statistically significant manner but mathematically only self-efficacy scores of EFL reading increased within the students in the control group from pre- to post-test.

Contrary to previous analyses, no effort to calculate *Cohen's d* was performed since no significant difference was detected in the mean scores of self-efficacy in EFL scale.

4.4. Statistics for Pre- and Post-test Mean Scores of the Experimental Group

This section of the analyses shed lights on the impact of flipped classroom intervention on EFL learners' academic achievement and its sub-sections, their attitudes toward and self-efficacy beliefs in EFL. In this respect, this part stands for one of the most important features of the current study. To enlighten the concerned bodies about the possible impact of the flipped classroom on EFL performance and its predictors, the researcher repeated the tests performed for control group. In other words, based on parametric structure of the data as previously revealed, separate *Paired Sample T Tests* were computed to explore whether students' means of EFL achievement test, attitude toward EFL and self-efficacy belief in EFL scales differ significantly from pre- to post-test within the experimental group.

Within this respect, following a similar way as in the presentation of the findings that answer the first research question of the study, the researcher demonstrated the

results of *Paired Sample T Tests* in three separate tables. Accordingly, the first of these tables shows the impact of flipped intervention on differences in the mean scores of whole achievement test and its sub-sections from pre- to post-test, the second table seeks its impact on the students' means of attitudes toward EFL scale within the control group. On the other hand, the third table finalizes the section by producing the results about the possible differences in the students' means of total scores of whole self-efficacy beliefs in EFL scale and its sub-scales within the experimental group.

Table 4.15: Paired Samples T-Tests for EFL Achievement Test

Factor	Group	N	$\overline{\mathrm{x}}$	S	$Sh_{\bar{x}}$	t	Df	p	d
EFL Listening	**Pre-test**	21	12,00	3,271	,714	-5,567	20	,000	-1,24
	Post-test	21	15,76	2,773	,605				
EFL Grammar	**Pre-test**	21	8,67	4,768	1,040	-8,662	20	,000	-2,13
	Post-test	21	16,57	2,158	,471				
EFL Reading	**Pre-test**	21	9,48	3,696	,807	-6,541	20	,000	-1,90
	Post-test	21	14,76	1,300	,284				
EFL Vocabulary	**Pre-test**	21	5,29	4,113	,897	-5,567	20	,000	-1,64
	Post-test	21	10,38	1,465	,320				
EFL Writing	**Pre-test**	21	11,10	5,691	1,242	-2,620	20	,016	-0,75
	Post-test	21	14,57	3,218	,702				
Total EFL Test Scores	**Pre-test**	21	46,52	17,503	3,820	-7,268	20	,000	-1,89
	Post-test	21	72,05	7,586	1,655				

As seen in Table 4.15, six separate Paired Samples t tests elaborate on the measures of EFL achievement test as a whole and in skills within the experimental group

where the students were taught based on a flipped classroom-teaching model for 15 weeks during the fall term. Therefore, this table informs the readers about the impact of the intervention on the students' EFL test from pre- to post-test. As a result, the table indicates the significant impact of flipped intervention on students' EFL test performance as a whole (t = -7,268; p <.001) and in all the sub-sections as listening (t = -5,567; p <.001), grammar (t = -8,662; p <.001), reading (t= -6,541; p <.001), vocabulary (t = -5,567; p <.001) and writing (t = -2,620; p <.05).

The second fold of these analyses presented the effect size of the aforementioned significant difference between EFL achievement pre- and post-test scores. Therefore, Cohen's *d* (degree of impact) was calculated to detail each of these significant differences. In this respect, based on Cohen's guidelines (Ruscio, 2008), it was interpreted that flipped intervention resulted in a large effect size of the significant difference in listening, grammar, reading, vocabulary and total EFL test achievement scores within the students in the experimental group (*d*>.80). However, *cohen's d* value suggested moderate effect size of significant difference in those students' scores of EFL writing (.50<*d*< .80). On the other hand, it is important to note that Cohen's (1994) suggestion of ignoring the positive or negative that are caused by positive or negative t value and taking the absolute value was again taken into account in the interpretation of *d* values.

In addition to the table on the means of the repeated measures regarding EFL achievement test and its five sub-sections within the students in the experimental group, the following table presents the findings that compare the difference in these students' mean scores of Attitudes toward EFL scale from pre- to post-test.

Table 4.16: Paired Samples T-Tests for Attitudes toward EFL

Factor	Groups	*N*	$\overline{\text{x}}$	S	$Sh_{\bar{x}}$	t-Test			
						t	*Df*	*P*	*d*
Attitudes Toward EFL	**Pre-test**	21	3,243	,2095	,0457	-2,115	20	,047	-0,69
	Post-test	21	3,431	,31911	,0696				

As seen in Table 4.16, *paired sample t test* yielded results that seek the possible difference in students' mean scores of Attitudes toward EFL scale within the experimental group from pre- to post-test. Accordingly, the table indicates the flipped classroom intervention differed the students' means of *attitudes toward EFL* significantly from pre- to post-test (t = 2,115; p <.05). However, on the contrary to negative significant impact on the scores of attitudes within the students in the control group, the difference in means developed positively from pre- (3,243) to post-test (3,431).

To detail the extent of the significant impact of the flipped intervention on the participant students' scores of Attitude toward EFL scale, *Cohen's d* was computed and evaluated based on Cohen's guidelines (1994). As a result of this, Cohen's d of -0.69 proved the moderate effect of the difference created by the flipped intervention in students' scores of attitudes from pre- to post-test administration (.50<*d*< .80).

Finally, the following table presents the results of five separate Paired Sample T tests to elaborate on the statistical difference in the mean scores of Self-efficacy beliefs in EFL scale within the students in the experimental group.

Table 4.17: Paired Samples T-Tests for Self-Efficacy Beliefs in EFL

Factor	Group	N	$\overline{x}$	S	$Sh_{\overline{x}}$	t-Test			
						t	Df	p	d
Self-Efficacy in EFL Reading	Pre-test	21	2,536	,7574	,1653	-6,339	20	,000	-1,33
	Post-test	21	3,530	,7362	,1607				
Self-Efficacy in EFL Writing	Pre-test	21	2,505	,4811	,1050	-3,279	20	,004	- 0,82
	Post-test	21	2,995	,6866	,1498				
Self-Efficacy in EFL Listening	Pre-test	21	2,676	,6707	,1464	-3,322	20	,003	0,49
	Post-test	21	3,014	,6887	,1503				
Self-Efficacy in EFL Speaking	Pre-test	21	2,667	,9250	,2018	-2,143	20	,045	- 0,55
	Post-test	21	3,143	,7858	,1715				
Total Self-Efficacy in EFL Scores	Pre-test	21	2,596	,6114	,1334	-4,464	20	,000	- 0,90
	Post-test	21	3,171	,6563	,1432				

As seen in Table 4.17, five separate *Paired Samples T tests* were run to explore the impact of flipped intervention on the mean scores of self-efficacy beliefs in EFL scale within the students in the experimental group from pre- to post-test administration. Within this respect, it was explored that flipped classroom instruction differed total mean scores of self-efficacy belief in EFL at $p < .001$ level (t = - 4,464). Moreover, this table also reveals the significant impact of the flipped instruction on students' beliefs in their own capacities to accomplish certain goals in such four main skills of EFL as reading (t= - 6,339; $p <.001$), writing (t= -3,279; $p <.01$), listening (t = - 3,332; $p <.01$) and speaking (t = -2,143; $p <.05$).

To better interpret the extent of the aforementioned significant impact of flipped classroom on EFL students' self-efficacy beliefs, *Cohen's d* was calculated for each of five significant values in table 4.13. Following Cohen's guidelines (Ruscio, 2008), it was interpreted that flipped intervention resulted in a large effect size of the significant difference between pre- and post-test scores of total self-efficacy beliefs

in EFL scale, EFL reading and writing within the students in the experimental group (*d*>.80). However, *Cohen's d* value suggested moderate effect size of significant difference in those students' scores of self-efficacy beliefs in EFL listening and speaking (.50<*d*< .80). On the other hand, it is important to note this interpretation of *d* values was based on Cohen's (1994) suggestion of ignoring the positive or negative that are caused by positive or negative t value and taking the absolute value.

To sum up, this section revealed the positive impact of flipped classroom intervention on EFL students' academic performance and its predictors in detail by elaborating on their degrees of impact through calculating *Cohen's d* value.

4.5. Statistics for Comparison of Post-test Mean Scores between the Groups

The current section of the data analyses stand for the distinctive feature of the study. In other words, it enlightens the concerned bodies about the comparative impact of flipped classroom instruction and traditional lecture-based teaching on EFL learners' academic achievement and its sub-sections, their attitudes toward and self-efficacy beliefs in EFL. To display the comparative impact of these two instructional designs on EFL students' post-test scores of *Achievement Test, Self-Efficacy Belief in* and *Attitudes toward EFL Scales*, *Independent Sample T Tests* were administered. The core results of the current study were rooted in the administration of these tests since factorial analysis, also known as MANOVA and Two-Way MANOVA, are mainly used to indicate either the effect of two independent variables on more than one or more dependent variables or that of a single independent variable on two or more dependent variables simultaneously covering the probable relationships among the variables avoiding type 1 error (Büyüköztürk, 2007). Within this respect, in the current study, the researcher equated the groups in terms of three variables as indicated in the aforementioned tables. In addition to that, the researcher does not aim to explore the joint impact of independent variables or the interaction between them on the dependent variable. Since the students in two groups were found to be equal in terms of their pre-test scores, the researcher only sought to explore the impact of the two types of instruction on the post-test scores of three data collection instruments. In other words, this aims to reveal the impact of instructional design standing for independent variable on EFL achievement, attitudes toward and self-efficacy beliefs in EFL standing for three different dependent variables of the study.

In this regard, based on the results that prove the data distributes normally, the researcher chose to employ separate *Independent Samples T Tests* to gain insights into the comparative impact of flipped classroom and traditional lecture-based teaching models on EFL achievement, attitudes toward and self-efficacy beliefs in EFL.

Within this respect, following a similar way to the one in the first two sections that focused on the significant differences in pre- and post-test scores within the groups, the researcher organized the presentation of the results of *Independent Samples T Tests* in three separate tables. Based on that, the first of these tables covers the results that search whether students in control and experimental groups differ significantly in terms of their post-test scores of whole EFL achievement test and its sub-sections. On the other hand, the second table investigates the impact of the instructional design on the students' post-test mean scores of attitudes toward EFL scale. Finally, the third of the tables aims to explore if the type of the instruction the students were provided exerts any significant impact on their post-test mean scores of whole self-efficacy beliefs in EFL scale and its sub-scales.

Table 4.18: Independent Samples T-Tests for EFL Achievement between the Groups

Factor	Group	N	$\overline{x}$	S	$Sh_{\overline{x}}$	t	Df	p	d
EFL Listening	Control	20	14,70	3,373	,754	-1,103	39	,277	----
	Exper.	21	15,76	2,773	,605				
EFL Grammar	Control	20	14,40	3,235	,723	-2,540	39	,015	-0,78
	Exper.	21	16,57	2,158	,471				
EFL Reading	Control	20	12,75	3,932	,879	-2,178	39	,040	-0,68
	Exper.	21	14,76	1,300	,284				
EFL Vocabulary	Control	20	8,10	4,412	,986	-2,200	39	,038	-0,69
	Exper.	21	10,38	1,465	,320				
EFL Writing	Control	20	12,20	4,047	,905	-2,082	39	,044	-0,65
	Exper.	21	14,57	3,218	,702				
Total EFL Test Scores	Control	20	62,15	15,513	3,469	-2,575	39	,016	-0,81
	Exper.	21	72,05	7,586	1,655				

Based on the results of the multiple analyses that set the data distribute normally, six separate *Independent Samples t tests* were run to explore if participant students' post-test mean scores of EFL achievement test differ significantly between control and experimental groups. In other words, Table 4.14 compared participant EFL students' academic performance for the key aspect of the study. In this respect, the findings in the table depict the participant students' quality of EFL performance under the impact of traditional lecture-based and flipped classroom instruction models. As a result, significant differences in the post-test mean scores of such sub-scales of EFL achievement test as EFL grammar (t= -2,540), EFL reading (t= -2,178), EFL vocabulary (t= -2,200), EFL writing (t= -2,082) were found between control and experimental groups (p <.05). On the other hand, another significant difference in the post-test mean scores of total EFL achievement test was detected between the groups (t = -2,575; p <.05). It is also important to state that all these significant differences,

as proved by the mean scores in the table, were found to be in the favour of experimental group where the learners were instructed following a flipped classroom model. However, the participant students' post-test scores of EFL listening did not differ significantly among groups (t = -1,103; p >.05).

The second fold of the current analyses was devoted to better enlighten the concerned bodies about the size of the significant impact of flipped classroom on EFL performance. With this purpose in mind, Cohen's *d* (degree of impact) was computed to delve into each of the significant differences stated above. First, it must be noted that the interpretation of objective effect size of the significant differences was realized based on three ranges provided by Cohen as "small = 0.2 < *d* < 0.5," "medium = 0.5 < *d* < 0.8,"," and "large = *d* >.80" (Ruscio, 2008). Also, Cohen's (1994) suggestion of ignoring the positive or negative that are caused by positive or negative t value and taking the absolute value shaped the ultimate interpretation. In this respect, all *Cohen's d* values in the table suggested moderate effect size of the significant difference created by flipped instruction in EFL grammar, reading, vocabulary and writing (*.50<d< .80*). Moreover, *Cohen's d* of -0,81 show flipped classroom instruction resulted in large effect size of significant difference in total EFL achievement test performance.

To sum up, the current analyses clearly explored the impact of flipped classroom instruction model on participant students' performance in total EFL achievement test and its sub-skills based on the *independent samples t test* enriched by *Cohen's d* calculation. Following this, next table seeks the difference in post-test mean scores of Attitudes toward EFL scale among control and experimental groups.

Table 4.19: Independent Samples T-Tests for Attitudes toward EFL between the Groups

Factor	Groups	*N*	$\overline{\text{x}}$	S	$Sh_{\bar{x}}$	t-Test			
						t	*Df*	*p*	*d*
Attitudes toward EFL	Control	20	2,90	,7382	,16507	-2,966	39	,006	0,90
	Exper.	21	3,431	,3191	,06963				

Independent samples t test results investigate the difference in students' post-test mean scores of *Attitudes toward EFL Scale* between the control and experimental groups. In other words, the table compares traditional and flipped classroom model for their impact on students' attitudes toward EFL. Within this respect, as seen in the table, the students in the flipped classroom were found to produce significantly better scores of attitudes toward EFL than those in the traditional classroom (t = -2,966; p <.01).

To standardize the significant impact of flipped classroom instruction on students' attitudes toward EFL objectively regardless of the variables, sample sizes or measurement types, the researcher referred to *Cohen's d*. Based on Cohen's guidelines (1994) to interpret *d* values, it was explored that the flipped classroom model exerted large size of significant difference on students' scores of attitudes toward EFL (*d*< .80).

In addition to these two tables above, the following table embodies the results that question the statistical difference in the post-test mean scores of Self-efficacy beliefs in EFL scale among control and experimental groups.

Table 4.20: Independent Samples T-Tests for Self-Efficacy Beliefs in EFL between the Groups

Factor	Group	N	$\bar{x}$	S	$Sh_{\bar{x}}$	t-Test t	Df	p	d
Self-Efficacy in EFL Reading	Control	20	2,925	,7951	,1778	-2,529	39	,016	-0,79
	Exper.	21	3,530	,7362	,1607				
Self-Efficacy in EFL Writing	Control	20	2,705	,6597	,1475	-1,379	39	,176	-----
	Exper.	21	2,995	,6866	,1498				
Self-Efficacy in EFL Listening	Control	20	2,875	,4423	,0989	-,774	39	,444	-----
	Exper.	21	3,014	,6887	,1503				
Self-Efficacy in EFL Speaking	Control	20	2,645	,6312	,1411	-2,230	39	,032	- 0,70
	Exper.	21	3,143	,7858	,1715				
Total Scores of Self-Efficacy in EFL	Control	20	2,788	,5289	,1183	-2,051	39	,047	- 0,64
	Exper.	21	3,171	,6563	,1432				

As seen in Table 4.20, five separate *Independent Samples T tests* were run to explore the impact of instructional design on post-test mean scores of self-efficacy beliefs in EFL scale among the students in control and experimental groups. In other words, the current part aims to exhibit the impact of flipped classroom instruction on students' beliefs in their own capacities to accomplish certain goals not only in EFL as a whole but also in its four sub-skills such as reading, writing, listening and speaking. Within this respect, significant difference in the total post-test scores of self-efficacy beliefs in EFL scale was detected among control and experimental groups (t = -2,051; p <.05). Additionally, it was also explored that the instructional design exerted a significant impact on participant students' post-test scores of two sub-scales of Self-efficacy Beliefs in EFL Scale: EFL reading (t = -2,529; p <.05) and EFL speaking (t = -2,230; p <.05). As understood from the related mean scores in the table, all these significant differences were in the favor of flipped classroom.

However, no significant difference in post-test scores of two self-efficacy beliefs in EFL writing and listening (t= -1,379, -,774; p>.05).

To specify the extent of the aforementioned significant impact of flipped classroom on EFL students' self-efficacy beliefs in an objective manner, the researcher calculated *Cohen's d* for each of three significant values in table 4.16. Following Cohen's guidelines to interpret the calculated values (Ruscio, 2008), it was noted that flipped classroom intervention moderately differed participant students' total scores of self-efficacy beliefs in EFL scale and its two sub-scales as EFL reading and speaking from pre- to post-test (.50<*d*< .80).

To sum up, this section was grounded in t tests results and *Cohen's d* values to comprehensively manifest the impact of flipped classroom intervention on EFL students' academic performance, learning attitudes and beliefs in their own capacities to succeed in EFL.

4.6. Statistics for Retention of EFL Achievement Test

The current section serves to expand the results that set the impact of the instructional design on EFL achievement performance. Hence, this part covers the attempts to explore the impact of the instructional design on the retention of students' performance in EFL achievement test. Furthermore, the researcher delves into this possible impact of the instruction design on sub-scales of EFL Achievement test.

The data for the current analyses was collected through the administration of EFL Achievement Test as a retention test two weeks after the post-test session. The interval time between the retention test and post-test was specified according to the pilot study. In the pilot study, two weeks after the post-test session, the participant students were asked to take the EFL achievement test again and reflect on the items they remembered. Based on the responses obtained, it was seen that most of the items (88 %) were not remembered by the participants. This indicated two-week interval was appropriate to adopt the EFL Achievement Test as an instrument of measuring the retention of EFL achievement.

In brief, the attempt to gain insights about the impact of flipped classroom on the retention of EFL achievement is embodied in the following results of six separate Independent Samples T tests.

Table 4.21: Independent Samples T-Tests for Retention of EFL Achievement

Factor	Group	N	$\overline{x}$	S	$Sh_{\overline{x}}$	t	Df	p	d
EFL Listening	Control	20	13,40	3,2831	,7341	-1,204	39	,236	----
	Exper.	21	14,52	2,6762	,5839				
EFL Grammar	Control	20	14,60	1,7290	,3866	-2,754	39	,009	-0,85
	Exper.	21	16,14	1,8516	,4041				
EFL Reading	Control	20	12,65	3,8289	,8561	-2,266	39	,034	-0,72
	Exper.	21	14,67	1,1106	,2423				
EFL Vocabulary	Control	20	9,25	1,5174	,3393	-2,929	39	,005	-0,92
	Exper.	21	10,86	1,9567	,4269				
EFL Writing	Control	20	12,75	3,4469	,7708	-2,311	39	,026	-0,72
	Exper.	21	14,95	2,6168	,5710				
Total EFL Test Scores	Control	20	62,65	9,9539	2,226	-3,374	39	,002	-1,05
	Exper.	21	71,14	5,6946	1,243				

Based on the results of the multiple analyses that set the data distribute normally, six separate *Independent Samples t tests* were executed to explore the impact of instructional designs on the retention of EFL performance between groups. Within this perspective, the figures in Table 4.17 clearly indicate the enhancing impact of the flipped classroom instruction on the retention of EFL students' academic achievement (p < .01; .05). In other words, compared to the students instructed following a traditional lecturing model, students in the flipped classroom were found to be significantly better at retaining their performance in total EFL achievement test (t= -3,374; p < .01). Yet, their significant superiority was also apparent in such sub-scales of the test as EFL grammar (t= -2,754; p < .01), EFL reading (t= -2,266; p < .05), EFL vocabulary (t= -2,929; p < .01) and EFL writing (t= -2,311; p < .05). On the other hand, Table 4.17 also reveals that type of the instructional design did not

exert any significant impact on the retention of students' performance in EFL listening (t= -2,311; p > .05).

In this respect, the findings in the table depict the participant students' quality of EFL performance under the impact of traditional lecture-based and flipped classroom instruction models. As a result, significant differences in the retention mean scores of such sub-scales of EFL achievement test as EFL grammar (t= -2,574), EFL reading (t= -2,266), EFL vocabulary (t= -2,929), EFL writing (t= -2,311) were found among control and experimental groups (p <.05). On the other hand, another significant difference in the mean scores of overall EFL achievement test was detected among the groups (t = -3,374; p <.05). It is also important to state that all these significant differences, as proved by the mean scores in the table, were found to be in the favor of experimental group where the learners were instructed following a flipped classroom model. However, the participant students' scores of EFL listening in the retention test did not differ significantly among groups (t = -1,204; p >.05).

The second fold of the current analyses set the size of the significant impact of flipped classroom on the retention of EFL achievement test performance in standard and objective terms. As in the previous sections, this was realized through calculating Cohen's *d* (degree of impact) for each of the significant differences stated above. It must be noted that ranges and guidelines provided by Cohen (1994) shaped the interpretation of the calculated *d* values. In this respect, taking the absolute value regardless of negative or positive, the flipped classroom instruction exerted moderately significant impact on the retention of EFL reading and writing test performance (.50< *d* < .80). On the other hand, flipping the classroom for EFL students was explored to have a large significant impact on the retention of performance in not only total EFL achievement test but also EFL grammar and vocabulary (*d*> .80).

In brief, the current analyses reflected on the quality of the impact of flipped classroom on EFL performance by exploring its role in the retention of performance in total EFL achievement and its sub-skill areas. Grounded in six separate *independent samples t tests* enriched by *Cohen's d* calculation, the size of the significant difference the flipped classroom instruction created in the retention of EFL performance was clearly displayed.

4.7. Qualitative Results

The present embedded study was performed in two phases. The first phase was grounded in quasi-experimental pre- and post-test design to explore the impact of the flipped intervention on EFL learners' academic performance and its predictors such as attitudes toward and self-efficacy beliefs in EFL. Within this respect, the data collected from quantitative sources was analyzed to indicate this probable impact through t-tests and variance analyses. However, the researcher organized the second phase in phenomenological design to gain insights about the experimental outcomes of the study. To facilitate the findings about quantitative results, the data collected from semi-controlled interviews with the participation of 9 experiment group students about their flipped classroom experiences was analyzed. These 9 students were specified by employing the *multiple variation sampling* to "present multiple perspectives of the individuals… that differ on some characteristic or trait" (Creswell, 2012, 208-209). In this respect, 9 experiment group students who developed differently from pre- to post- academic achievement test were interviewed. Therefore, 3 of these students were chosen from those whose scores were the most negatively or least positively developed while 3 other students were those who moderately changed. On the other hand, the remaining 3 interviewee students were those whose scores were the most positively changed from pre to post-test. Each of the interviews was held in the university in a period of 25-40 minutes. During this period, the researcher asked the participants ten questions, formed after the review of the related literature, about their flipped classroom experiences, their attitudes toward and self-efficacy beliefs in EFL. Based on the participants' consent, these interviews were realized and the researcher audiotaped by the researcher.

The collected data were analyzed qualitatively by employing content analysis method. There seems to be no accepted steps to follow to analyze qualitative data. Yet, such an analysis included "developing a general sense of the data, and then coding description and themes (also called category) about the central phenomenon" (Creswell, 2012, 237). Therefore, in the current study, the analysis started with transcription of the recorded data. Following it, the transcribed data was manually analyzed. Based on Creswell's recommendations, the transcribed data was coded. Then, to provide the accuracy of the codes, the researcher asked the interviewee to reflect on these codes. These codes, formed with the consensus of the participants,

were evaluated based on 9 themes. These themes were pre-determined by the researcher excluding the last question since it was an open-ended one to allow the interviewee to elaborate on their flipped experience in any way. This ongoing process of forming codes that sometimes lead to wider codes continued till the saturation. After it was ensured that no codes would arise and all the codes and themes were verified, their frequencies were calculated using SPSS 21 for Windows.

The researcher also attempted to achieve the reliability and accuracy of qualitative results including the themes established based on the interview questions, and the codes formed based on interviewee's responses. However, reliability and validity are addressed in different terms such as *trustworthiness, rigor, quality, dependability* and ensured in different ways in qualitative studies (Golafshani, 2003). Therefore, the researcher followed some common approaches reviewed in the literature to establish validity and reliability of qualitative findings (Creswell, Miller, 2000; Golafshani, 2003; Guba, Licoln, 1982; Yıldırım and Şimşek, 2006). First of all, as set by Patton (1990), the researcher himself is the instrument in the qualitative research. Therefore, the researcher of the study thickly described the setting, every step of qualitative data collection and analysis procedures to attain the credibility of the researcher's interpretation. On the other hand, to further validate the accuracy of the researcher's interpretations, the researcher received feedback from the interviewee and confirmed the codes formed out of their responses by also achieving the fairness of the interpretation. This process, also called as member checking (Creswell, 2012; Golafshani, 2003), did not alter any of the codes since the participant students found the researcher's interpretation appropriate. In terms of the dependability or consistency of qualitative analysis, the researcher used two approaches. First, through triangulation, the researcher drew on the information obtained from both surveys and the interviews to judge the quality of the results. Moreover, the researcher also employed *peer debriefing* method which is defined by Creswell (2012) as a process during which the researcher works with impartial colleagues or experts. They examine collected data, codes, themes and reports and provide necessary feedback. Based on these feedbacks, the researcher employed Miles and Huberman's (1994) formula to document this consistency. Miles and Huberman (1994) set the final inter-coder agreement rate in qualitative data analysis as 70% or above. Based on one colleague's and one expert's feedbacks, the stated formula

where Reliability = agreement / [agreement + disagreement] X 100) was performed. As a result, the pre-determined themes and codes were found appropriate with an agreement of 92% regarding the themes and 89% regarding the codes.

The results of the qualitative analysis were presented in three folds. The reason behind it was three-dimensional structure of the interview. In this respect, the first dimension included first six questions in the interview to reflect on the impact of flipped classroom on EFL achievement. In other words, first five of the questions in the interview aimed to explore participants' perceptions of the role of flipped classroom in the main components of their EFL performance while the last one focused on the expansion of it into other courses. On the other hand, the second fold included two questions about the impact of flipped classroom on their attitudes toward EFL. The ninth question formed the third fold of the analysis by seeking how flipped classroom influenced EFL students' beliefs in their capacity to succeed in EFL. The last question was used to enrich the responses to the other questions.

Within this perspective, the following table summarizes the results of qualitative analysis regarding the first six questions in the interview.

Table 4.22: Categories and Codes for the perception of Flipped Classroom in EFL

Categories	Codes			
Advantages of Flipped Classroom	Varying teaching materials	Preparation for class	Easy to learn	Systematic approach
N	1	4	3	1
%	11,1 %	44,5 %	33,3 %	11,1 %
Disadvantages of Flipped Classroom	Technical		Unfamiliar	
N	8		1	
%	88,9%		11,1 %	
Impact of Flipped Classroom on EFL Learning Style	Studying regularly		Active participation	
N	7		2	
%	77,8 %		12,2 %	
Impact of Flipped Classroom on time to study out of school	Positive		Neutral	
N	8		1	
%	88,9 %		11,1 %	
Flipping EFL classrooms	Positive			
N	9			
%	100 %			
Flipping other classrooms	Positive		Doubtful	
N	6		3	
%	66,7%		33,3 %	

The first six questions in the semi-structured interview form questioned students' perceptions of flipping EFL classroom and how it affected the way they study English. As seen in the table, these questions details the impact of flipped EFL classroom on academic performance by delving into its advantages and disadvantages, how it differed students' learning style and the time they spent out of

school to learn EFL. Within this respect, the table indicates that most of the students, no matter how well they performed during the term, have positive attitudes toward flipping EFL classroom. Specifically, nearly half of the students (44,5 %) stressed that flipped classroom helped them a lot come to school ready for the class. As an example, they express this advantage of the instruction stating

> *"Thanks to the presentations I watched and quizzes I took, I always went to school having studied the subject"* (Student 1)

In addition to 1 student (11,1 %) who saw the positive side of flipped classroom instruction as the use of *"a rich array of teaching materials"* (Student 4), another student stated that it made him *"study regularly and systematically since"* he *"knew what and how to come next"* (Student 6). On the other hand, 3 other students (33,3 %) regarded flipped classroom as an instruction that made learning fun and easy by stating

> *"By strengthening the relation between students and teacher, it makes us learn faster in a more comfortable atmosphere."* (Student 3)

In terms of the disadvantages, except for 1 student who said *"I have never been involved in such an instruction"* (Student 3), all the other students (88,9 %) indicated technical problems such as poor network connection or smart devices as a drawback of flipped classroom by stating

> *"I had some small problems. They are all, of course, about the internet and computer. Sometimes, I could not see the assigned videos. Sometimes it was difficult to take the quizzes on my smart phone when there is no computer."* (Student 9)

Students were explored to think that flipped classroom changed their learning behaviors in two different ways. Firstly, most of them (77,8 %) pointed out flipping presentation and homework routine contributed to their learning stating

> *"It made me study regularly."* (Student 5)

> *"Before I was studying once or twice for my exams. I was doing this for two or three days before the exam. I was getting tired a lot. Thanks to this instruction, studying for 15-20 minutes a day, I learnt everything about the content and I didn't need to study extra for my exam."* (Student 7)

226

On the other hand, 2 students (12,2 %) thought it helped them actively participate in the lesson since they had an idea about what they would learn in the classroom. They stated

> *"By assigning continuous and little homework, it does not bore the students and foster student participation in the lesson."* (Student 4)

The fourth question aimed to gain insights about how much time the students spent on EFL learning out of the school during the term when their course was flipped. As seen in the table, students' responses were coded as positive and neutral. On the contrary to 1 student who stated *"I can't say that I spend much time except for the homework."*, 8 students (88,9 %) had positive ideas about flipped classroom thinking that it helped them manage their time better. Different ideas were expressed as in the following ways:

> *"It made me share time to study before I sleep."* (Student 7)

> *"Since it helped me manage my time better and use it more productively, I learned a lot in a shorter time. It 50 to 60 % decreased the time I spent out of school for the lesson."* (Student 6)

> *"I learn better and I spend less out of school."* (Student 4)

All the students (100 %) reported that they support the idea of flipping EFL classrooms. They justified themselves in such similar ways as

> *"It made me learn better"* (Student 7)

> *"It must be used to facilitate student participation."* (Student 2)

> *"I learned better by spending less time. Thanks to quizzes, I saw my mistakes and focused on them more."* (Student 8)

In terms of flipping other courses, most of the students (66, 7 %) seemed to be hesitantly positive. They usually stated their opinions in a way similar to the following:

> *"I would like it to be used in other courses as well. But I am not sure if it works as well as in English."* (Student 1)

"Yes I want. Because having idea about the course before the classroom makes it easier and funnier to learn". (Student 7)

However, 3 of the students (33,3 %) seemed to be hesitant about flipping courses other than EFL stating

"It may be useful for some verbal courses. But I am not sure about science and math classes." (Student 9)

"I would like but it may be illogical for all the courses." (Student 5)

As seen in the table, these first six questions details the impact of flipped EFL classroom on academic performance by delving into its advantages and disadvantages, how it differed students' learning style and the time they spent out of school to learn EFL. Based on the results in Table 4.18, it was better understood that flipped classroom model positively triggered university prep EFL students' academic performance and increased the time they spent to learn and study.

In addition to these qualitative results supporting the related quantitative results, the following table serves to foster the understanding about the impact of flipped classroom model on attitudes toward EFL as a predictor of EFL performance.

Table 4.23: Categories and Codes for the Impact of Flipped Classroom on Attitudes toward EFL

Categories	Codes	
Attitudes toward EFL	Positive	
N	9	
%	100 %	
Teaching and learning atmosphere	Learning with fun	Easy to learn
N	8	1
%	88,9%	11,1 %

The seventh and eighth questions in the semi-structured interview form sought the impact of flipped classroom instruction on EFL students' attitudes toward EFL course from their perspectives. In this respect, students from different achievement

groups ranging from low to high were asked to elaborate on how flipped classroom affected their attitudes toward EFL and how flipping course presentation and assignments differed teaching and learning atmosphere in the classroom. Within this respect, "attitudes toward EFL" and "teaching and learning atmosphere" themes were formed. As seen in the table, flipped classroom instruction exerted a positive impact on all students' attitudes toward EFL. All the students based this positive change on its systematic approach to teach that makes it easier for them to learn. This is clearly certain in the following example responses given by students from three achievement groups:

> *"Flipped classroom removed my negative ideas about EFL."* (Student 1)

> *"I think going to the classroom ready thanks to presentations sent through Edmodo and extra quizzes contributed to my language development"* (Student 3)

> *"I am not a person who is into English. But this instruction changed my perception about EFL since it made it easier to learn English."* (Student 6)

> *"Normally I am not a student interested in English courses. However, this instruction forced me to participate in the courses and helped me learn."* (Student 5)

When it comes to how online assignments and higher order in-class activities differed teaching and learning atmosphere in EFL classroom, most of the students (88,9 %) stressed fun element in the flipped EFL course. They also seemed to justify their ideas based on flipping structure of the courses. These students indicated that teaching the content at home and help them practice on it through online quizzes reinforced their learning. These all made them find the course funnier and easier to learn. This was expressed by one of the interviewed students as in the following:

> *"It made learning funny. Think. Do you feel good when you talk about something that you know something about or something that you don't know anything about? This instruction makes learning funnier and more comfortable."* (Student 6)

However, only one of the students (11,1 %) did not mention anything about fun element of flipped classroom. Instead, the student stressed how easy it was to learn with flipped classroom by stating,

"Instead of trying to learn the content, we learned by reinforcing. We did not lose time on unnecessary details and focused on our mistakes. We did more exercises both individually and in groups. This all made learning English easy." (Student 8)

As a result of the responses given by nine students to seventh and eighth questions in the interview, all the students, no matter how well they performed during the term, had positive attitudes toward flipping EFL classroom. It stressed it was easy and fun to learn English through flipped classroom instruction.

The following table, on the other hand, presents the coded qualitative analysis to enrich the results about the impact of flipped classroom instruction on EFL students' self-efficacy beliefs.

Table 4.24: Categories and Codes for the Impact of Flipped Classroom on Self-efficacy Beliefs in EFL

Categories	Codes	
Self-efficacy beliefs in EFL	Positive	Neutral
N	8	1
%	88,9%	11,1 %

The ninth semi-controlled interview question aimed to explore EFL students perceptions about the role of flipped classroom instruction in another reviewed predictor of EFL achievement, which is self-efficacy beliefs in their talents to succeed in EFL. In this step, students from different achievement groups ranging from low to high were asked to give the general sense of the impact of flipped classroom on their self-trust in their ability to perform overall in EFL without going into details about four separate skills. The researcher, as seen in the table, coded the students' responses as "positive" and "neutral" under "self-efficacy beliefs in EFL" theme.

As seen in Table 4.20, it can be stated that the qualitative results are parallel to the related quantitative findings. In other words, the table indicated none of the interviewees thought flipped classroom instruction made them lose their self-

confidence in their ability to succeed in EFL. As seen in the table, eight of the students (88,9 %) reported that flipping EFL classroom encouraged them to be better EFL learners. On the contrary to those who simply said "*Yes, it increased my self-confidence.*", other students were in line with each other. They stated and justified this empowering impact of flipped classroom instruction on their self-confidence as in the following ways:

> "*Actually, it showed me that English was not that difficult and in this way helped me gain self-confidence.*" (Student 4)

> "*Since I learned the content much more easily than I thought, it affected my self-confidence and belief in a positive way.*" (Student 1)

However, as seen in the table, one of the students (11,1 %) did not believe in empowering impact of flipped classroom on self-confidence in the ability to succeed in EFL. The student simply stated her response as follows:

> "*It had no effect on my self-belief in being successful in EFL.*" (Student 9)

As a result of the analysis of the response given to the ninth interview question, nearly all the students from different achievement groups were found to be in the favor of flipped classroom since they thought it triggered their self-efficacy beliefs in being successful EFL learners.

In addition to nine semi-controlled questions in the interview, the last tenth question let them elaborate on flipped classroom in any way they want. Only two of the students were eager to answer the last question. They expressed their opinions about flipped classroom instruction in a way that strengthen the qualitative analyses reached in this section. They elaborated on flipped classroom instruction as follows:

> "*Flipped classroom instruction is one of the most important and effective model that must be followed in the classroom not only for students but also for teachers.*" (Student 4)

> "*Although possible technical problems form the drawbacks of the model, I believe that it is very useful for students' learning and development*" (Student 7)

To sum up, the qualitative analysis in this part sought to deepen the understanding of the impact of flipped classroom instruction on EFL performance and attitudes toward

and self-efficacy beliefs in EFL as the predictors of the achievement in it. As a result of the analyses of semi-controlled interviews with the participation of 9 experiment group students about their flipped classroom experiences, the acquired qualitative results facilitated experimental outcomes of the study. In other words, confirming the results acquired from quantitative sources through t-tests and variance analyses, the qualitative results enriched the perception of the efficiency of flipped classroom.

To conclude, the qualitative results helped the researcher gain insights about such advantages of flipped classroom instruction as increasing the time students spend to prepare for the class and actively participate in the class, development of higher-level skills. Last but not the least, the results that explore correlation with more positive attitude toward and stronger self-efficacy beliefs in EFL attract the attention to flipped classroom instruction as a facilitator of EFL performance while they also increase the awareness of some technical problems that may hinder its efficiency.

5. CONCLUSIONS AND IMPLICATIONS

This embedded study grounded in a mixed method research design adopted flipped classroom model in teaching EFL courses of a prep class at Gebze Technical University for the whole process of 2016-2017 fall term. Adopting a traditional lecture based teaching approach in another classroom, the researcher aimed to examine the effects of flipped classroom model on not only EFL students' academic performance but also their attitudes toward and self-efficacy beliefs in EFL. The necessary data to achieve this aim was collected through implementation of EFL achievement test, Attitudes toward EFL and Self-efficacy Beliefs in EFL scales as pre and post-test. The results were enriched through semi-controlled interviews with nine students from three achievement groups in flipped EFL classroom.

The current chapter of the study draws conclusions about the statistical findings in a way to stand for the answers of the research questions respectively. In this vein, the results of the present study are compared with those of the similar studies reviewed in the relevant literature. From this point, the chapter seeks to provide answers to the research questions by integrating and synthesizing the issues raised in the whole study.

On the other hand, the current chapter serves to explore the implications of flipped classroom model in EFL. In this respect, the first fold covers theoretical, practical and policy implications for practicing EFL teachers and curriculum designers. The second fold of the implications, on the other hand, presents direction to prospective researchers for future research. Eventually, by elaborating on all aspects of the issue, the chapter aims to carry all concerned bodies to a new level of understanding about flipped classroom in EFL.

5.1. Conclusions

Dwelling on the quantitative and qualitative analyses run within the scope of the study, the conclusions of the study are elaborately discussed to answer the research questions respectively.

5.1.1. Research Question 1: Is there a statistically significant difference between pre- and post-test scores of the students in the traditional lecture-based classroom with regard to

a) EFL Achievement Test,
b) Sub-sections of EFL Achievement test
c) Attitudes toward EFL Scale,
d) Self-efficacy Beliefs in EFL Scale,
e) Sub-scales of Self-efficacy Beliefs in EFL?

The main focus of the study was to seek how flipped classroom instruction differed EFL students' academic performance, attitudes toward and self-efficacy beliefs in EFL from pre- to post-test administration. However, the researcher also sought the effects of traditional lecturing on the variables of the study. Within this respect, the first research question examined the development of these students in terms of their EFL performance, their attitudes toward and self-efficacy beliefs in EFL within control group.

In the first fold, the related results of the study, in this respect, indicated that traditional lecturing positively differed students' total EFL performance from pre- to post-test with a moderate effect size. Moreover, this type of instruction was also found to have impact on their performance in listening, reading, grammar, and vocabulary with a moderate to large effect size. Yet, it did not significantly affect their EFL writing performance. This increase in EFL performance should not be thought that traditional lecturing is superior to flipped classroom model. Any kind of instruction can be expected to increase students' academic performance compared to their performance before the instruction.

The reasoning behind this increase in EFL performance through traditional instruction seems to be supported by the results of some other studies conducted with similar designs in EFL. Similarly, Leis, Cooke and Tohei (2015) compared development of composition skills of EFL Japanese students in traditional and flipped classrooms. Since the students in two groups were significantly different at the beginning of the study, the researchers compared the changes within each group. This makes the results of the study relevant to the current study. Consistent with the present study, Leis, Cooke and Tohei (2015) explored the students in the traditional

group were able to produce significantly higher mean scores with a medium effect size. This result is also in parallel to the result of another PhD thesis conducted by Ekmekçi (2014) flipping a writing course for EFL students in School of Foreign Languages at Ondokuz Mayıs University. As a result of pre- and post-test true experimental design study, Ekmekçi (2014) compared writing performance of the students both between and within groups. Within group comparison indicated that traditional lecturing also positively and significantly differed students' EFL writing performance.

In addition to the consistency with the relevant results about positive impact of traditional lecturing on EFL performance within itself, insignificant difference in writing performance of the students in the control group from pre- to post testing is mostly because writing is the hardest skill for EFL learners. This can be justified by Chamot (2005) who attracts the attention to how difficult it is for EFL students to write in EFL regardless of their levels. Chamot (2005) elaborates on this by stating that "beginning level students struggle with finding the words they need and remembering grammatical conventions, whereas more advanced students find it difficult to link their ideas with coherence and to produce appropriate target language discourse" (121).

The second fold of the first research question examined how students' perception toward EFL differed as a result of traditional lecturing. The study indicated that traditional lecturing had a negative impact on EFL students' attitudes toward EFL with a medium effect size. A bulk of studies in the literature that shows a strong relationship between attitudes and language learning performance (Fakeye, 2010; Gardner, Lambert, 1972; Gardner, 1985; Gömleksiz, 2010; İlhan, Karatas, 2015; Şentürk, 2015; Tsiplakides, Keramida, 2010; Rukh, 2014). Yet, this negative attitude toward EFL, regardless of academic performance, can also be partly explained by students' age groups since students are thought to hold more negative attitudes toward EFL as they get older (Cenoz, 2001). However, it is nearly impossible to crosscheck these results of the present study because there seems to be no study to seek attitudes toward traditional EFL classrooms. These studies seem to be grounded in experimental designs comparing the attitudes between groups not within them. Yet, Lasagabaster and Sierra (2009) seem to support the attitudinal decline observed in traditional EFL classrooms. The researchers (2009) compared the attitudes

students enrolled in the CLIL (Content and Language Integrated Learning) groups hold toward EFL and those in traditional EFL groups do. Consistent with the present study, Lasagabaster and Sierra (2009) explored that students in traditional EFL classroom have less positive attitudes toward EFL.

The third and last fold of the first research question focuses on the impact of traditional EFL classroom on their self-efficacy beliefs in EFL. The present study explored that traditional lecture-based instruction model had no significant impact on neither self-efficacy beliefs in EFL as a whole nor four sub-skill areas as reading, writing, listening, and speaking. Yet, it is important to state that the means of self-efficacy in EFL writing, listening, speaking and total scores decreased from pre- to post-test administration of the scale. Similar to the reasoning about the lack of studies to crosscheck the results about attitudes toward EFL, it is again not that possible to find studies in the related literature based on self-efficacy beliefs in EFL. Most of such studies were experimentally designed to compare the impact of the intervention between groups. In other words, in EFL, "research studies have examined self-efficacy in relation to a limited number of variables namely learning strategies, performance, causal attributions, and language anxiety" (Raoofi, Tan, Chan, 2012, 61). Despite this, as explored by İlhan and Karataş (2015), self-efficacy beliefs, also termed as motivational beliefs, are significantly correlated to attitudes toward EFL and they are also indicated as strong predictors of academic performance in EFL (Rahimi, Abedini, 2009; Cotteral, 1999). In this respect, negative attitude toward EFL among the students in the traditional classroom may trigger lack of beliefs in their ability to succeed in EFL. This insignificant impact of traditional lecture-based model on self-efficacy in EFL can be linked some other affective factors such as lack of motivation (İlhan, Karataş, 2015), high level of anxiety (Erkan, Saban, 2011), lack of strategy training or use (Khajavi, Ketabi, 2012) and some contextual factors such as past experiences (Çakır, Alıcı, 2009), role of teacher (Moghari et al., 2011) and classroom climate (Greta, 2009).

To conclude, despite lack of studies in the literature to crosscheck the related results of the current study, the first research question stressed the decline in attitudes toward and self-efficacy beliefs in EFL. As significantly correlated with each other, attitudes and self-efficacy beliefs were set as the strong predictors of academic performance in EFL. However, the traditional instruction still exerted a positive

impact on EFL performance. This decline in attitudes and self-efficacy beliefs may be more apparent in the comparison between groups.

5.1.2. Research Question 2: Is there a statistically significant difference between pre- and post-test scores of the students in the flipped classroom with regard to

a) EFL Achievement Test,

b) Sub-sections of EFL Achievement test,

c) Attitudes toward EFL Scale,

d) Self-efficacy Beliefs in EFL Scale and,

e) Sub-scales of Self-efficacy Beliefs in EFL?

The second research question of the study aimed to explore the effect of flipped classroom intervention on the development of three dependent variables of the study from pre- to post-test administration. Within this respect, in this phase of the study, the researcher attempted to shed lights on how flipped classroom instruction differed the students in the experimental group in terms of their EFL performance, their attitudes toward and self-efficacy beliefs in EFL from pre- to post-test administration of the quantitative data collection instruments.

Repeating the similar tests performed for control group, the researcher elaborated on the impact of the flipped classroom on EFL performance and its predictors. As the first step of the second research question, the results pointed out significant impact of flipped classroom intervention on total EFL performance and all its sub-skills with a large effect size. It was found that only students' EFL writing performance was moderately differed by flipped classroom instruction within experimental group from pre- to post-test administration. This moderate effect size of the intervention on EFL writing can be explained by Chamot's (2005) stress on it as the most difficult skill to improve among EFL students regardless of their levels of language.

The facilitating impact of flipped classroom model on EFL performance is supported by a number of studies across STEM subject areas across K-12 and higher education contexts such as science (Bergman, Sams, 2012; Reid, 2016), technology (Davies, Dean, Ball, 2013), math (Fulton, 2013; Johnson, 2013), in Language Arts (Fulton, 2013) and pharmacy courses (Ryan, 2013). On the contrary to this heavy reliance on the use of flipped classroom in STEM disciplines in generally K-12 contexts, "little or no research to date has rigorously studied whether and how flipping the language

classroom can enhance student learning" (Hung, 2015, 83). Furthermore, there seems to be a lack of efficient number of studies searching the effect of the flipped instruction on sub-skills in EFL. Yet, this limited related literature also provides consistent results with those of the current study that indicate the positive impact of flipped classroom on EFL students' performance within experimental classroom. Engin (2014, 21) is one of a number of researchers who state that research in the related literature "explored the effectiveness of the flipped classroom model". Farah (2014) also proved this positive impact of the instruction on EFL writing. In a master study conducted in United Arab Emirates K-12 context, the flipped classroom instruction resulted in significant improvement in writing performance within experimental group. In another study conducted with adult learners in an ESL course at a community college in US, Han (2015) explored that adopting a flipped approach improved students' both oral fluency and grammar explanations. Similar to them, Leis, Cooke and Tohei (2015) showed facilitating impact of flipped classroom on EFL writing performance within experimental group in a Japanese higher education context. Another quasi-experimental mixed method study was conducted by Webb and Doman (2016) in two universities in the US and China. The results indicated that grammar scores in the post-test were significantly higher than those in the pre-test within experimental group. Hung (2015), on the other hand, compared students' EFL performance in three different classrooms termed as *structured, semi-structured* and *non-flipped classrooms*. The results indicated that students in structured flipped classrooms outperformed those in other two classrooms.

To the researcher's scope of interest, only 2 of 11 studies in Turkey conducted about flipped classroom were directly related to EFL. In this respect, the present research was thought to enrich the related literature through its results gathered from its mixed method design. As one of those two studies, Ekmekçi (2014), as a result of his doctorate study conducted in a higher education context in Turkey, supported the significant impact of flipped classroom model on EFL writing exploring significant progress in students' EFL writing performance from pre- to post-test administration within experimental group. In the second and the last study conducted in Turkish higher education context, Boyraz (2014) indicated that the flipped classroom exerted a significant impact on EFL grammar performance within experimental group with a large effect size. Although grammar covered only two structures as "reported

speech" and "passives" in that study, there seems to be a perfect match between the results of that study and the related ones of the current study. As for achievement dimension, only one of the studies on the flipped classroom in various disciplines, reviewed by Bishop and Verleger (2013, 12), was reported to "examine student performance throughout a semester". By examining the role of flipped classroom in EFL for 15 weeks during the whole fall term, the current study fills an important gap in the literature. Furthermore, the study is proved to be in line with the general tendency in the related literature that shows triggering impact of flipped classroom on academic performance across different disciplines in different contexts. This can simply be attributed to the fact that "technology is the important support to develop the students learning" (Kitchakarn, 2015, 52) in today's world where technology has become an indispensable part of educational settings.

In addition to the consistency with the relevant results about positive impact of flipping EFL classroom on total EFL performance, the study explored its significant enhancing role in all sub-skill areas in EFL. Due to the lack of necessary studies to crosscheck the related results of the study, the studies were limited to those searching grammar, writing and overall EFL performance. Therefore, this aspect of the present study is expected to fill in this gap in the literature.

The second fold of the second research question examined students' attitudes toward flipping EFL classroom within experimental group. The results of the study indicated that students' attitudes toward EFL differed significantly in a positive manner with a moderate effect size from pre- to post-test administration. Consistent with what was reported by Webb and Doman (2016), most of the studies on students' perceptions of and attitudes toward flipped classroom explored that students are overwhelmingly positive toward it. In line with them, Bishop and Verleger (2013) elaborate on positive attitudes toward flipped classroom in various disciplines as follows:

"Despite differences among studies, general reports of student perceptions were relatively consistent. Opinions tended to be positive, but there were invariably a few students who strongly disliked the change" (10).

In addition to a number of studies showing positive attitudes toward flipped classroom across different subject areas (Bergman, Sam, 2012; Davies, Dean, Ball, 2013; Long et al., 2016; Love et al., 2014; Pedroni, Meyer, 2006; Wilson, 2013;

Zappe et al. 2009), it is a must to evaluate the related results of the study through similar studies in EFL. Although they are not many in number, they seem to share the same basis with the tendency in the general literature (Al-Harbi, Alshumaimeri, 2016; Hung, 2015; Webb, Doman, 2016). In similar quasi-experimental designs, these studies yielded positive attitudinal findings toward flipping EFL classroom in different contexts through the responses to surveys and interviews. Overlapping with this body of international studies, the current study proves itself in terms of the students' positive attitudes toward flipped learning environments.

To the researcher's scope of interest that matches with that of the study, there seems to be only three studies conducted about how students perceived flipped classroom instruction in EFL in higher education contexts in Turkey (Basal, 2015; Boyraz, 2014; Ekmekçi, 2014). Directly related to the current study, through surveys followed by interviews or open-ended questions analyzed thematically, Basal (2015) Boyraz (2014) and Ekmekçi (2014) revealed that students hold positive attitudes toward and made positive comments about flipped classroom in EFL. Revealing the results in the favor of positive attitudinal impact of flipped classroom similar to those in the literature, the current study set flipped classroom as a model that enhances motivation and creates positive attitudes toward EFL by increasing extrinsic, intrinsic motivation and helping students better manage their cognitive load (Abeysekera, Dawson, 2015).

The last fold of the second research question examines how flipped classroom instruction differs students' self-efficacy beliefs in EFL within experimental group from pre- to post-test administration. *Paired Samples T tests* revealed that flipped classroom instruction significantly differed students' total self-efficacy beliefs in EFL in a positive manner with a large effect size. The results also indicated the similar impact of those students' self-efficacy beliefs in four skill areas with a moderate to large effect size. However, to the researcher's scope, the literature seems to lack of such a study integrating flipped classroom instruction and self-efficacy beliefs in EFL. As also set by Raoofi, Tan and Chan (2013), in EFL context, self-efficacy is largely studied in terms of its relation to "a limited number of variables namely learning strategies, performance, causal attributions, and language anxiety" (61). However, most of these studies seem to examine the relation between self-efficacy beliefs and performance in one of the sub-skills in the target language in

different contexts (Abedini, Rahimi 2009; Doordinejad, Afshar, 2014; Erkan, Saban, 2011; Khajavi, Ketabi, 2012; Osman et al., 2016; Tilfarlioğlu, Cinkara, 2009; Todaka; 2013). Within this perspective, the related results of the study serve to enrich the literature on self-efficacy in EFL by relating it to flipped classroom in EFL. Despite the lack of studies to back this unique aspect of the study, the study seems to be in line with the literature on self-efficacy in EFL where it is assigned to be the predictor of learners' motivation and academic performance. This significant impact of the flipped classroom model on self-efficacy in EFL can be attributed to the personalized learning and teaching atmosphere in flipped classroom. In this atmosphere, students learn at their own paces, better involved with engaging activities, their attention are attracted and beliefs in their abilities to succeed are enhanced through this self-directed learning (Basal, 2015; Hung, 2015; Yu, Zhu, 2016; Zappe et al., 2009). When the significant relation between self-efficacy beliefs in and attitudes toward EFL is taken into consideration, the chain-relation is easily understood that flipping EFL classroom forms positive attitude and that improves self-efficacy as a result of which the process produces satisfied students.

To conclude, lack of studies to crosscheck the results of the current study about the impact of flipped classroom on self-efficacy beliefs in EFL clearly indicated valuable contribution of the present study to the literature on flipped classroom in EFL. On the other hand, the study yielded results similar to most of the studies in the literature. In this respect, it set flipped classroom as a facilitator of EFL performance, positive attitudes toward and self-efficacy beliefs in EFL by also dwelling on the significant correlation between attitudes and self-efficacy as the predictors of that performance.

5.1.3. Research Question 3: Is there a statistically significant difference between post-test scores of the students in the flipped classroom and traditional lecture-based classroom with regard to

> **a)** EFL Achievement Test,
> **b)** Sub-sections of EFL Achievement test
> **c)** Attitudes toward EFL Scale,
> **d)** Self-efficacy Beliefs in EFL Scale and,
> **e)** Sub-scales of Self-efficacy Beliefs in EFL?

The main purpose of this study was to gain insights about the efficacy of flipped classroom instruction in improving EFL performance and two most cited predictors of it as attitudes and self-efficacy. Dwelling on the limitations of a large number of studies carried out on this issue, the present study suggested valuable results for the related literature by seeking the impact of flipping EFL classroom on enhancing performance in all sub-skills and two affective factors as its predictors in a relatively longer period that lasted a whole term. In this respect, the third research question of the study was devoted to shed lights on the comparative impact of flipped classroom instruction and traditional lecture-based teaching on EFL learners' academic achievement and its sub-sections, their attitudes toward and self-efficacy beliefs in EFL. In this respect, Independent Sample T Tests were administered on the two groups of students' post-test scores.

Different from the first two research questions that focus on the role of flipped classroom in EFL within groups, the third research question formed the core of the current study organized in a distinguished design to elaborate on all aspects of flipping EFL classrooms between groups in a single study. However, the related literature includes a large number of isolated studies that focus on one of these three main aspects of the study. In other words, some of these studies deal with the effect of flipped classroom model on improving EFL performance in not all but some of sub-skills in international (Engin, 2014; Farah, 2014; Leis, Cooke and Tohei, 2015; Hung, 2015; Webb and Doman, 2016) and national contexts (Boyraz, 2014; Ekmekçi, 2014). On the other hand, some other studies were limited to explore how this model, compared to traditional lecturing, differed students' attitudes toward EFL in international (Al-Harbi, Alshumaimeri, 2016; Hung, 2015; Webb, Doman, 2016) and national contexts (Basal, 2015; Boyraz; 2014; Ekmekçi, 2014). However, there seems to be no study that directly relates self-efficacy to flipping EFL classrooms despite a number of studies that adopt it as a predictor of performance in different skills in a target language (Abedini, Rahimi 2009; Doordinejad, Afshar, 2014; Erkan, Saban, 2011; Khajavi, Ketabi, 2012; Osman et al., 2016; Tilfarlioğlu, Cinkara, 2009; Todaka; 2013). Although these studies do not hold three variables within the interest of the study as a whole in a single study, they still yield isolated results to crosscheck with those of the current study.

The first fold of the third research question aimed to enlighten the concerned bodies about the comparative efficiency of flipped classroom model in EFL performance in different skills, which stands for the distinguished aspect of the study in the related literature. As stated earlier, different from the second research question that sought the impact of flipped classroom on EFL performance from pre- to post-test administration within the students in the experimental group, this question compared the difference in EFL post-test performance between flipped and traditional classrooms. As a result of running six separate Independent Samples T Tests, students' EFL performance in all skill areas except for listening was significantly under the impact of flipped classroom instruction model. Students in the flipped classroom were found to perform significantly better than those in the traditional lecture based classroom. While the effect size of the difference in grammar, reading, vocabulary and writing was found to be moderate, that in overall EFL performance was explored to be large. These results of the current study seem to expand the related literature by applying flipped classroom to non-STEM disciplines. In addition to producing similar results to those studies grounded in non-STEM disciplines (Davies, Dean, Ball, 2013; Kim et al., 2014; Love, Hodge, Grandgenett, Swift, 2014; Mclaughlin et al., 2013; Moravec et al., 2010; Reid, 2016; Strayer, 2007; Zappe, et al., 2009), the present study produced results supported by insufficient number of studies conducted in international context (Ahmad, 2016; Al-Harbi, Alshumaimeri, 2016; Hung, 2015; Kang, 2015; Roth, Suppasetseree, 2016; Webb, Doman, 2016). Webb and Doman (2016) also supported contribution of flipped classroom instruction to EFL learning outcomes. The researchers investigated the impact of flipped classroom instruction on EFL grammar performance in ESL and EFL contexts in the US and China. As a result of this experimental study, students in the flipped classroom were found to have significant difference in gains on grammar compared to those students in non-flipped classrooms. Kang (2015) also produced consistent results with the current study in a Korean EFL context by exploring that the flipped classroom was significantly more effective than traditional lecturing in improving students' grammar, vocabulary and overall English performance. Contrary to this and current study, Al-Harbi and Alshumaimeri (2016) explored that flipped classroom did not exert a significant difference in students' grammar performance in Saudi Arabian K-12 context. Yet, it is important to note that although the difference was not statistically significant, students in the flipped classroom still outperformed

243

those in the traditional classroom. Al-Harbi and Alshumaimeri (2016) attribute this insignificant difference to some "unique challenges of the Saudi context" (89), On the other hand, Hung (2015), who aimed to explore the impact of flipping EFL classroom on performance, attitudes and participation levels in a broader perspective, also proved the superiority of flipped classroom to semi-flipped and non-flipped classrooms in facilitating EFL performance. In a post-test-only quasi-experimental design, Hung (2015) designed classrooms in three different instructional formats: structured, semi-structured and non-flipped classrooms. As a result of assessing students' performance through a lesson assessment that was given in the second week of each lesson and included "vocabulary quizzes (30%), multiple choice questions for video comprehension (30%), and an oral presentation or writing performance (40%)" (Hung, 2015, 87), the students in the structured (totally flipped) classroom were found to perform significantly better than those in semi-structured and non-flipped classrooms. In terms of enhancing EFL students' listening abilities, the present study produced contrasting results with those of few studies in the literature (Ahmad, 2016; Roth, Suppasetseree, 2016). In Cambodian and Egyptian contexts, the researchers adopted one-group pre-test post-test designed study. Despite lack of opportunity to compare the effect of flipped classroom instruction to that of traditional classroom on EFL listening due to the mismatch between the designs of the present and mentioned studies, flipped classroom was found to significantly enhance university EFL students' listening scores from pre- to post-test. This contrast in the results can stem from the difference in the designs of the studies or participants' lack of practice in EFL listening since EFL listening includes many challenges for a EFL student, which can only be dealt with more and more practice.

On the contrary to these studies in the international context, there are fewer studies conducted on the comparative impact of flipped classroom on EFL performance in Turkish context (Boyraz, 2014; Ekmekçi, 2014). Conducted in a similar setting to that of the current study, Boyraz (2014) and Ekmekçi (2014) investigated the efficiency of flipped classroom in EFL grammar and writing respectively. Both of these studies dealt with flipped classroom in EFL higher education context. While Boyraz (2014) organized his study to include two structures in grammar that lasted for a few weeks, Ekmekçi (2014) applied the model during the whole fall term in 2013-2014 academic year. Despite some of their limitations, these studies also

validated the results of the present study by statistically proving the superiority of flipped classroom to traditional lecture based classroom in fostering EFL performance. Although it does not share the same basis as the current study, it is also necessary to evaluate one of the recent studies conducted by Kurt (2017) to expand the related literature in Turkey. Kurt (2017) designed a pre-test post-test quasi-experimental study to compare the roles of flipped and traditional lecture-based classrooms in improving self-efficacy beliefs and learning outcomes in a classroom management course at a university in Turkey. Although it did not produce results directly related to EFL, Kurt (2017) explored that flipped classroom resulted in significantly better performance in final exam than traditional classroom.

As a valuable contribution to not only national but also international literature, the present study compared flipped classroom instruction to traditional lecture-based instruction to explore its effect on students' performance in overall EFL achievement test and the sub-scales of it. As a result, the study was enriched by confirming its fostering impact on overall EFL, grammar, writing and vocabulary performance explored by limited number of studies focusing on some of the sub-skills. On the other hand, the study pioneered to expand the literature by showing its impact on EFL reading and insignificant effect on EFL listening. This determinant impact of flipped classroom on EFL performance can "stem from the self-pacing opportunity in the learning environment" (Kurt, 2017, 218). Student-centered and more personalized learning atmosphere in the classroom, flipping instructional design helped the learners to be engaged in higher-levels of Bloom taxonomy, study on the video content regardless of time and place through their own pace (Uzunboylu, Karagozlu, 2015, Zappe et al., 2009).

The second fold of the third research question compared flipped classroom instruction to traditional lecture-based design in order to explore how it affects students' attitudes toward EFL. Within this respect, Independent samples t test results revealed that students in the flipped classroom held significantly more positive attitudes toward EFL than those in the traditional classroom with a large effect size. This large determinant impact of flipped classroom on positive attitudes toward EFL was found to be consistent with the positive attitudinal findings about flipped learning in the related literature (Ekmekçi, 2014). These positive attitudes toward flipped classroom instruction are also stated to exist in various disciplines

other than EFL (Bergman, Sam, 2012; Bishop, Verleger; 2013; Davies, Dean, Ball, 2013; Long et al., 2016; Love et al., 2014; Pedroni, Meyer, 2006; Wilson, 2013; Zappe et al. 2009). However, there is a "limited amount of research currently available on flipping, while also exploring the viability of flipping at the tertiary level with non-native English speaking students" (Webb, Doman, Pusey, 2014). Despite some minor differences in the study designs and lack of direct comparison with traditional classroom, the present study is also in line with the general positive attitudes toward flipped classroom in EFL in international (Ahmad, 2016; Al-Harbi, Alshumaimeri, 2016; Hung, 2015; Roth, Suppasetseree, 2016; Kang, 2015; Webb, Doman, 2016) and national context (Basal, 2015; Boyraz, 2014; Ekmekçi, 2014). In other words, following pre-test post-test or only post-test quasi-experimental designs, these studies indicated that students' perceptions about flipping EFL classrooms were overwhelmingly positive. This positive attitude toward flipping EFL classroom can be linked to its "potential to cater for motivation and cognitive load" (Abeysekera, Dawson, 2015, 10). In other words, coming to the classroom ready for the course, using various e-learning and online materials or resources available and accessible independently of time and place differed significantly students' attitudes toward EFL by engaging them more in the course (Basal, 2015, Kostka and Lockwood, 2015). Furthermore, positive attitude toward flipping EFL can also be attributed to the fact that flipped classroom allows students to share their work with their friends and teachers, get feedback from them at any place and time in addition to reducing their anxiety by donating them with opportunity to think on their responses without being under pressure to give an immediate answer (Al-Harbi, Alshumaimeri, 2016).

On the other hand, distinguished than those studies in EFL or other STEM disciplines, the current study does not measure students' attitudes toward flipping EFL through administering surveys or semi-structured interviews. The study, as a response to the second fold of the third research question, aimed to compare flipped classroom instruction to traditional classroom in a pre-test post-test quasi-experimental design to determine their impact on attitudes toward not flipping EFL classroom but EFL itself. In this respect, though partly related to those cited studies, this aspect of the study speaks for itself and promises valuable information to be taken into account by concerned bodies. The reason behind this is the close

relationship between positive attitudes toward EFL and achievement in it since motivation includes learners' perceptions of the content determining the amount of academic effort during the process of learning the language (Yurtseven, Alcı, Karataş, 2014).

The last fold of the third research question stood for another pioneering aspect of the current study. This fold was devoted to compare the effect of flipped EFL classroom and traditional lecture-based classroom on EFL learners' self-efficacy beliefs. This study detailed this contribution to the related literature by seeking this impact on four sub-scales of self-efficacy beliefs in EFL. As a result of five separate Independent Samples T tests, the flipped classroom instruction was found to exert a significant positive impact on not only EFL students' overall self-efficacy beliefs in EFL but also those in EFL reading and speaking with a moderate effect size. Despite its significant impact on all sub-scales of self-efficacy within experimental group, it is interesting to see its superiority to traditional classroom in only self-efficacy in EFL reading, speaking and total scores of the scale. This can be attributed to cited significant correlation between self-efficacy beliefs and attitudes toward EFL (İlhan and Karataş, 2015). In other words, the positive attitudes can better help EFL students in the flipped classroom to deal with their EFL speaking anxiety since flipped classroom can be a reaction to reasons why Turkish EFL students feel anxious to speak in EFL. In other words, it can clarify some of the reasons listed by Demir (2015) behind speaking anxiety in EFL such as "speed at which lessons develop" "feeling unprotected when giving spontaneous" and "peer reactions" (42). On the other hand, the fact that flipping EFL classroom help the learners to reinforce grammatical structures better than traditional classroom through online quizzes, short exams and permanent control of the content might enhance self-efficacy in EFL reading.

This unique aspect of the study that integrates flipped classroom and self-efficacy in EFL gains much more importance when the strength of self-efficacy as an affective factor playing an important role in EFL performance is understood (Yanar, Bümen, 2012). In addition to a large number of studies that set self-efficacy as a strong predictor of academic achievement (Rahimi, Abedini, 2009), similar to attitudes, it is determinative on one's amount of effort performed to succeed in something (Bandura, 1997). Therefore, self-efficacy in EFL is largely studied in a relation to

some affective factors, learning strategies or performance in one of the sub-skills in the target language in different contexts (Abedini, Rahimi 2009; Doordinejad, Afshar, 2014; Erkan, Saban, 2011; Khajavi, Ketabi, 2012; Osman et al., 2016; Raoofi, Tan, Chan, 2013; Tilfarlioğlu, Cinkara, 2009; Todaka; 2013). Despite this bulk of studies showing self-efficacy as a strong predictor of EFL achievement, there seems to be no effort in related literature to relate it to flipped classroom. Within this perspective, this facilitating impact of the flipped classroom on self-efficacy beliefs in overall EFL, reading and speaking in EFL over traditional classroom cannot be crosschecked. Despite that, correlation between attitudes and self-efficacy as strong predictors of academic performance in EFL (Rahimi, Abedini, 2009; Cotteral, 1999) can validate the related results. In other words, significant decline in the scores of attitudes and statistical but not significant decline in scores of self-efficacy within control group can better explain why flipped classroom instruction was superior to traditional lecture-based classroom in enhancing self-efficacy beliefs in EFL. Furthermore, significant positive difference in the self-efficacy scores within the flipped classroom further increased the gap between groups in terms of self-efficacy in EFL. On the other hand, features of flipped classroom that tap on self-directed learning and individualized learning attract the learners' attention by embedding the course to technology. Moreover, it fosters their beliefs in their abilities to succeed in EFL by allowing them to learn the content at their own pace independently of time and place (Basal, 2015; Hung, 2015; Yu, Zhu, 2016; Zappe et al., 2009).

To conclude, the third research question is unique in terms of expanding growing literature about the flipped classroom model. To researcher's scope, to date, there are no studies in the related literature adopting flipped classroom model together with self-efficacy and attitudes in a comparison to traditional lecture-based classroom in a single study grounded in a pre-test post-test quasi-experimental design. By pioneering to integrate strong predictors of EFL achievement in Turkey to flipped classroom approach, the related results of the study can be informative for the curriculum designers and instructors concerned with EFL at universities. In this respect, flipped classroom can be a remedy to foster EFL students' academic performance in higher education institutions by varying their' learning experiences, meeting their expectations and tapping on their powers of affective individuality in a

collaborative and blended learning context where the design of the classroom is reversed.

5.1.4. Research Question 4: Is there a statistically significant difference in the retention of EFL achievement test performance between the students in the flipped classroom and traditional lecture-based classroom?

The fourth research question forms another distinguishing aspect of the research by comparing the impact of flipped classroom instruction on retention of overall performance in EFL and its sub-skills to that of traditional lecture-based classroom. To the researcher's scope, the related literature includes few studies seeking the long-term impact of the flipped classroom model on EFL performance. As a result of the lack of efficient number of studies, these results will be more meaningful when they are supported by future research.

The result for the current research question was acquired through six separate *Independent Samples T tests* after administering EFL Achievement Test as a retention test two weeks after the post-test session. Deciding on the interval time between the retention test and post-test based on the responses obtained from students, the results indicated flipped classroom was significantly more effective than traditional lecturing in long-term retention of performance in overall EFL and its sub-skills except for listening with effect sizes ranging from moderate to large. Although this result speaks for itself, Boyraz (2014) supports the present study by showing that flipped classroom was significantly better than traditional lecturing at retention of EFL performance in his study where students were to study on two grammar structures for a relatively short period. In addition to Boyraz (2014) whose study was limited to grammar in Turkish context, this significant impact of flipped classroom instruction in EFL can also be attributed to "increased contribution to the educational environment" (Uzunboylu, Karagözlü, 146). Chan (2009, 153) states that "retrieval can enhance long-term retention of tested material" and this is performed through "the level of integration invoked during encoding and the length of delay between retrieval practice and final test". Dwelling on these two factors as the facilitators of retention in EFL, it is easier to adopt flipped classroom as the predictors of these two facilitators. In addition to helping students achieve the content at all levels of Bloom taxonomy (Basal, 2015; Bergmann, Sams, 2012),

flipped classroom serves for these two variables by allowing students to learn class content outside the classroom and apply what they learn in the classroom where they collaborate with their peers and get immediate feedback from their teachers (Zappe et al., 2009). Flipped classroom instruction can also help learners keep their performance in a significantly long term by increasing class time for more engaging instruction (Milman, 2014). Addressing to all the students' learning styles through integrating the technology they are already familiar with outside into their classrooms (Kim et al., 2014). Based on all these cited benefits of the flipped classroom that integrate the technological world with which the students are already familiar into their classroom, the flipped classroom seems to enhance both affective factors such as attitudes and self-efficacy and intellectual dynamics such as integration and delay that lead to long-term permanence of EFL performance.

To conclude, in addition to a handful of research studies following similar designs that explored significant impact on student outcomes in different subject areas (Chao, Chen, Chuang, 2015; Day, Foley, 2006; Koo et al., 2016; Love, Hodge, Grandgenett, Swift; 2014; Moravec, 2010), the present study provided valuable results to expand the related literature by noting flipped classroom as a better facilitator of retention of academic performance in EFL than traditional lecturing.

5.1.5. Research Question 5: What are the EFL learners' perceptions of their learning experiences in the flipped classroom?

The major purpose of the current study was to inform a wide range of shareholders in EFL at higher education institutions ranging from teachers to curriculum designers and policy makers about a more qualified EFL practice through flipping instructional phases thanks to the advantages of technology. Taking cited advantages of flipped classroom model in the related literature into account (Ahmad, 2016; Al-Harbi, Alshumaimeri, 2016; Basal, 2015; Boyraz, 2014; Ekmekçi, 2014; Hung, 2015; Kurt, 2017; Yu, Zhu, 2016), the last two research questions of the present study were formed to facilitate quantitative outcomes of the study through implementing a semi-structured interview form developed by the researcher (see Appendix 11) to nine students from the flipped EFL classroom. To better understand the issue and come up with "information rich" data (Patton, 1990), the concluding two questions aimed to gain insights about participant EFL students' self-reported opinions about their

flipped classroom experience and its impact on two predictors of their EFL performance such as attitudes and self-efficacy beliefs. Using content analysis method (Creswell, 2012; Patton, 1990), the researcher transcribed the data, coded it according to the themes pre-determined by him to stand for each of the questions in the interview, classified the codes based on the themes and calculated the frequencies and percentages.

The results of the qualitative analysis of the present study were three-folded because of three-dimensional structure of the interview. In this respect, the first dimension stands for the seventh research question by aiming to elaborate on EFL students' perceptions of flipped classroom and its role on their academic performance in EFL based on the first six questions in the interview form. This research question of the study does not promise to come up with unique results to expand the existing literature. However, it aims to increase the validity and generalizability of the results that explore learners' positive views of flipped classroom model within various subject areas (Ahmad, 2016; Boyraz, 2014; Chen, Yang, Hsiao; 2015; Ekmekçi, 2014; Gilboy, Heinerichs, Pazzaglia, 2015; Hung, 2015; Roth, Suppasetseree; 2016, Strayer, 2009; Turan, 2015). In other words, tracing students' perceptions of flipped classroom through implementing flipped classroom attitude surveys or interviews following quantitative procedures is a frequently viewed feature of flipped classroom literature on both EFL and other disciplines. This is proved by the assertion in Hamdan et al. (2013) that "a modest amount of research exists from individual educators who practice the Flipped Learning model and their views on behalf of their pupils" (12). Hence, the present study served for strengthening the qualitative results about flipped classroom approach by exploring positive perceptions of flipping EFL classrooms among EFL students from different achievement groups. These results of the study confirmed the findings reported in Hung (2015) that design of the materials, enhancing interaction with teachers and classmates satisfied students to learn in a funny atmosphere. Consistent with Kurt (2017), the present study indicated that learners in the flipped classroom regarded "learning in such an environment helped them gain first-hand experience of the flipped classroom approach" and "sitting passively and listening to a lecture boring" (218). Being in line with Basal (2015), the study provides convincing results that flipping EFL classrooms promote students to preview the learning materials that trigger their participation in the

classroom. Consistent with Ekmekçi (2014), the present study indicates that students regard technical problems such as internet connection as the principal disadvantage of flipped classroom approach. Despite students' positive views on flipping EFL courses, they were explored to be hesitantly positive about the expansion of this model into other subject areas. This hesitance seems to be consistence with the results of Reid (2016) that set no statistically positive difference in the evaluation of flipped chemistry classroom. Moreover, the hesitance about the application of flipped classroom to other courses is also witnessed in the study conducted by Zappe et al. (2009) where the students in an architectural engineering course in a higher education context were found to "prefer that only about half the classes be flipped and some use of traditional lectures should be maintained" (13).

To sum up, grounded in a phenomenological design, the study took a further step to enlighten the concerned shareholders about how to maximize the potentials of the flipped classroom model. Overlapping with studies from different subject areas including EFL, the study explored that students overwhelmingly held positive views about flipped classroom (Bishop, Verleger, 2013). Confirming to the results of various studies in EFL in national and international scope (Ahmad, 2016; Basal, 2015; Boyraz, 2014; Ekmekçi, 2014; Hung, 2015; Kang, 2015; Kurt, 2017; Roth, Suppasetseree), the researcher explored that these students justified themselves through its such features that create opportunities to improve their active learning, engagement and satisfaction with the course by helping them learn at their own pace in a way to stop and rewind the instructional content independent of place and time. Finally yet importantly, despite some technical problems that can be easily eliminated by a careful planning and preparation, flipped classroom was regarded by students worth applying since it enhances their academic performance in and engagement with EFL.

5.1.6. Research Question 6: What are the EFL learners' perceptions of the impact of their learning experiences in the flipped classroom on their

    a) attitudes toward EFL and

    b) self-efficacy beliefs in EFL?

The last research question of the study was grounded in the aim to better enlighten the concerned bodies about the role of flipped classroom instruction model through

providing "information rich" data (Patton, 1990). This data was acquired through semi-structured interviews held with nine students from different achievement groups from the flipped EFL classroom. In this respect, following content analysis method (Creswell, 2012; Patton, 1990), the researcher attempted to come up with results to facilitate quantitative outcomes of the study.

Three-dimensional structure of the interview led the researcher to present the results of the qualitative analysis of the present study in three folds. In this respect, the first dimension sought to explore EFL students' self-reported perceptions of flipped classroom and its role on their academic performance in EFL through the first six questions in the interview form. Embodied in the seventh research question of the present study, the researcher solely aimed to serve the related literature to generalize the results that explore learners' positive views of flipped classroom model within various subject areas (Ahmad, 2016; Boyraz, 2014; Chen, Yang, Hsiao; 2015; Ekmekçi, 2014; Gilboy, Heinerichs, Pazzaglia, 2015; Hung, 2015; Roth, Suppasetseree; 2016, Turan, 2015).

On the other hand, the last research question of the study was grounded in the remaining two folds of the interview. Based on the remaining questions in the interview form from seven to nine, the researcher promised unique results that can expand the limits of the existing literature. This incomparable aspect of this current research question is clear since students' perceptions of their flipped classroom experience and its impact on their academic performance is overwhelmingly studied through flipped classroom attitude surveys or interviews following quantitative procedures in flipped classroom literature on both EFL and other disciplines (Ahmad, 2016; Basal, 2015; Chen, Yang, Hsiao; 2015; Gilboy, Heinerichs, Pazzaglia, 2015; Johnson; 2013; Koo et al.; Kurt, 2017; Love, Hodge, Grandgenett, Swift, 2014; McLaughlin et al.; 2013; Reid, 2016). Hamdan et al. (2013) also asserts "a modest amount of research exists from individual educators who practice the Flipped Learning model and their views on behalf of their pupils" (12). However, the current research questions aims to explore self-reported perceptions of the students in the flipped classroom regarding its impact on their attitudes toward and self-efficacy beliefs in EFL. These results promise attractive results since e great deal of importance has been attached to foreign language learners' attitudes, perception of learning and teaching, motivation and beliefs in their capacity to succeed as key

factors predicting achievement (Fakeye, 2010; Gardner, 1985; Gardner, Lambert, 1972; Rukh, 2014; Osman et al, 2016; Şentürk, 2015; Todaka, 2013). Similarly, a great deal of effort has been performed by the researchers in national context to relate successful and unsuccessful EFL performance in Turkey to attitudes and self-efficacy beliefs (Akkuş, 2009; Aküzel, 2006; Başaran, Cabaroğlu, 2015; Çatal, 2015; Gömleksiz, 2010; Yanar, Bümen, 2012; İlhan, Karatas, 2015; Kabaharnup, 2010; Kanadlı, Bağçeci, 2015; Kazazoğlu, 2013; Kiziltepe, 2000; Özen, 1979). By showing the significant correlation between the strength of positive attitudes and self-efficacy beliefs and EFL performance, the related literature has set attitudes and self-efficacy as vigorous predictors of EFL performance in both national and international context. Delving into this significant relation, the researcher aimed to trigger the efficiency of the flipped classroom model by seeking its impact on these two strong facilitators of EFL performance.

Unique qualitative aspect of the current research question speaks for itself and this makes it difficult to crosscheck its results with those of others in the related literature. In this respect, the seventh and eighth questions in the semi-structured interview form sought the impact of flipped classroom instruction on EFL students' attitudes toward EFL course. It was seen that all the students stated that flipped classroom positively affected not only their attitudes toward their EFL course by associating this to its facilitating impact on learning and teaching atmosphere. Although these qualitative results improve and support the insights about the related quantitative findings, it seems impossible to compare it with similar qualitative results in EFL. This is grounded in the fact that the related literature has dealt with students' perceptions of and attitudes toward flipping their classroom (Al-Harbi, Alshumaimeri, 2016). Yet, this positive impact of the flipped classroom on their attitudes toward EFL can be attributed to the fact that it makes them better understand the content and the availability of a number of online materials and resources embedded with technology and its potential to tap on EFL students' motivation (Abeysekera, Dawson, 2015; Obari, Lambacher, 2015). In this respect, despite contradicting with some few studies that report students' complains about excessive out of class assignments and lack of technology (Han, 2015; Kang, 2015), the results of the current study are satisfactorily supported by the literature (Ahmad,

2016; Al-Harbi, Alshumaimeri, 2016; Basal, 2015; Boyraz, 2014; Ekmekçi, 2014; Hung, 2015; Roth, Suppasetseree, 2016; Webb, Doman, 2016).

The last dimension of the qualitative results was based on the ninth question in the interview form to provide insights about the impact of flipped classroom on self-efficacy beliefs in their capacity to succeed in EFL. As another strongly cited predictor of successful EFL performance, the students were asked to elaborate on how their self-efficacy beliefs changed as a result of flipped classroom instruction. Being in line with quantitative results, the students reported to be in favour of empowering impact of flipped classroom on their trust in themselves to succeed in EFL. Despite the lack of studies that can be directly compared to the results of the current study, some features of flipped classroom such as helping EFL students come to class prepared and learn the content at their own pace independent of time and place are thought to tap on their beliefs in their talents to succeed in EFL (Basal, 2015; Hung, 2015; Yu, Zhu, 2016; Zappe et al., 2009).

The remaining tenth question of the interview, on the other hand, aimed to provide the interviewee with a chance to elaborate on flipped classroom model in any way they wanted. As a sum of all these interview questions and the responses provided by the students, the researcher promised valuable results to the wide range of concerned bodies while shaping and implementing the EFL curriculum at universities in Turkey. Mixed method embedded design of the study raised its importance since review of the related literature indicated that nearly all data gathered about the effect of flipped classroom on attitudes, self-efficacy and academic performance involved the administration of surveys and few of them examined pre- and post-test scores (Zainuddin, Halili, 2016). In this respect, enlightening the concerned bodies about the flipped classroom model through information rich data, the current study has set that students' perceptions of flipped classroom regarding its influence on their academic performance, their attitudes toward not the flipped classroom but EFL and self-efficacy beliefs in their ability to succeed in EFL are overwhelmingly positive. In other words, students' attitudes toward and self-efficacy beliefs in EFL were significantly promoted after being involved in the flipping EFL classroom compared to those in the traditional classroom.

Last but not the least, compatible to the general trend in flipped classroom research, the current study explored flipped classroom model as a significant facilitator of EFL

performance and its two vigorous predictors, attitudes toward and self-efficacy beliefs in EFL in Turkey. This impact was found to be exerted thanks to "action based, authentic, connected and collaborative, innovative, high level, engaging, experience based, project based, inquiry based, and self-actualizing" activities flipped classroom model provides in the classroom (Hamdan et al., 2013, 17) and its feature "allowing students to control the pace or stream of learning content" (Ibrahim, Callaway, 2014, 21). Considering all these, the present study, adopting a mixed method design, committed to delve into the important dynamics of EFL performance in Turkey. In this respect, associating attitude and self-efficacy as strong predictors of EFL performance with flipped classroom model, the present study aimed to propose flipping EFL classrooms as an efficient way of dealing with failure in EFL in Turkey by better responding the needs of students as digital natives.

5.2. Implications

In the present study, the researcher has discussed how flipped classroom instruction, a model shaped by the principles of flipped classroom model rooted in constructivism and blended learning, affected self-efficacy beliefs, attitudes and academic achievement among EFL students in higher education institutions in Turkey. In this respect, the present study promises valuable outcomes that guide the instructors and curriculum designers to come up with a new pedagogical model to enhance the quality of EFL teaching and learning processes at universities in Turkish context.

On the other hand, the flipped classroom's commonly cited superiority to the traditional approach signals a new trend in student learning and makes it inevitable to further study this hot call for pedagogical shift. Within this respect, based on the limitations of current study, the researcher also urges the future researchers to take the further steps to delve into this issue.

To conclude, this section is designed in two folds. While the first fold covers the pedagogical implications curriculum developers and teachers to deal effectively with the major obstacles faced by EFL students at prep schools of universities in Turkey, the second fold touches on the implications for future research to better understand all aspects of this facilitating new trend in education.

5.2.1. Pedagogical Implications

The analysis of the data collected about the role of flipped classroom in self-efficacy, attitude and EFL performance compared to that of traditional teaching approach led the researcher to propose flipped classroom model as one of the cure for the problems of EFL at universities in Turkey. Rooted in the principles of constructivism, blended learning and differentiated instruction that helps learners take control of their own learning in an individualized learning environment, the study pioneers to clearly depict planning and implementation processes of flipping EFL classrooms at prep schools in Turkish context.

Within commonly cited advantages and disadvantages of flipped classroom in the literature, the researcher takes the results of the present study into account. As a result, following suggestions are made for the instructors, curriculum designers and policy makers to ensure a qualified flipped EFL classroom:

✓ The findings of this study clearly indicated the significant effect of flipped classroom model on EFL achievement and self-efficacy and attitudes as its predictors. The results also set its success over traditional teaching approach. Therefore, these results in line with those in the literature raise important issues the implementation of flipped classroom model in EFL. In this respect, standing for the first and second pillar of flipped classroom that stand for "f" and "l" of F-L-I-P, teachers must embed their classes with technology both for implementation efficiencies and motivational factors. However, such integration requires comprehensive planning of the materials to be used in and out of the classroom activities with a focus on active learning. Such a planning of flipping sequences of in the classroom and outside the classroom procedures assigns the instructors with the followings tasks:

  • Flipped classroom model leads to a shift from teacher-centered learning atmosphere to a student-centered approach where students "can theoretically pace their learning by reviewing content outside the group learning space, and teachers can maximize the use of face-to-face classroom interactions to check for and ensure student understanding and synthesis of the material" (Hamdan, et al., 2013, 5).This student centered learning culture of flipped learning loads the teachers with crucial task to

adopt the students to this differentiated instruction shaped with the use of a great deal of media that leads to keeping a balance between direct instruction and scaffolding (Bergmann, Sams, 2012). This process of adaptation also requires teachers to inform their students about their shifting roles during the instruction, learning and teaching procedures including how they will be instructed and use course management system.

- First exposure to learning content outside the classroom is the crucial element of reversed procedures in flipped classroom instruction. This exposure is realized through videos, screencasts, and narrated PowerPoint presentations shared through a course management system. However, teachers are advised to keep those videos in 10 to 15 minutes ideally (Bergmann, Sams, 2012) since the related literature reported that students preferred shorter and live in-person video lectures more (Bishop, Verleger, 2013).

- Teachers must take Nielsen's (2012) warning into consideration and be careful not to steal students' time that they can spend socializing with their friends and families. However, they need to form a balanced approach to ensure that students watch the videos about the lecture content. This can be realized by assigning students to submit their reflections on each of those shared videos, presentation or materials through completing short quizzes or answering the questions about the lecture content before they come to their schools (Johnson, 2013).

- To ensure that students do necessary preparation for the class by watching the videos sent through a course management system, teachers are proposed to adopt an assignment-model where students are briefly asked about the lecture content through items in different types (Brame, 2013). Delving into students' responses to the items in online short quizzes following each video or presentation content, teachers are advised to give more productive feedback decreasing the heavily reliance on the lecturer. In this way, teachers will create "an opportunity for students to use their new factual knowledge while they have access to immediate feedback from their peers and the instructor … and help them learn to correct

misconceptions and organize their new knowledge such that it is more accessible for future use" (Brame, 2013, 3).

- Teachers of flipped classrooms are responsible for realizing "f" of flipped classroom to create a flexible environment by physically rearranging "their learning space to accommodate the lesson or unit, which might involve group work, independent study, research, performance, and evaluation" (Hamdan et al., 2013, 4). Such a concept of flexibility leads itself to flexibility in the teachers' "expectations of student timelines for learning and how students are assessed" (Hamdan, et al., 2013, 5).In this way, the students are ensured to take more control of their learning, which will trigger their beliefs in their abilities to succeed.

- Teachers must allocate most of the classroom time to enhance higher order skills among students by providing" individualized support as students work through the activities designed to help them master the material" (Hamdan et al., 2013, 4). Within this perspective, professional teachers in the flipped classrooms are those who "continually observe their students, provide them with feedback relevant in the moment, and continuously assess their work by being reflective in their practice, connecting them with each other to improve their trade, accept constructive criticism, and tolerate controlled classroom chaos" (Hamdan, et al., 2013, 6) to foster the development of higher order skills (Strayer, 2012).

✓ Curriculum designers and policy makers also need to take these results into account and design a qualified flipped curriculum for EFL at prep schools in Turkish higher education institutions. In other words, policy makers must realize the "f" of F-L-I-P that stands for the first pillar of flipped learning by creating a flexible environment. In this respect, policy makers must ensure that the infrastructure and network systems are suitable for the implementation of this model. In addition to their responsibilities to install necessary infrastructure in the classrooms, the policy makers are also given the following tasks to enhance the efficiency of the model:

- Through in-service education, policy makers must ensure "p" of F-L-I-P by training instructors to efficiently "utilize the affordances of the model to help students gain conceptual understanding, as well as procedural fluency when needed" (Hamdan et al., 2013, 6). To foster their efficiency, instructors can also be trained about how to use such course management systems as *Edmodo, Schoology* or *Moodle*, short video podcasts and screen casting technologies to record their screen or voices on those screens.

- To serve for purpose of using intentional content that stands for "i" of F-L-I-P, instructors must also be provided with a bulk of materials including pre-recorded videos enriched with various Web 2.0 tools, PowerPoint presentations voiced or narrated and appropriate short exams. Such an understanding will result in tailoring the content to support the curriculum in a collaborative and active learning atmosphere to facilitate construction, application and synthesis of the content knowledge (Bergmann, Sams, 2012). All these, in turn, will help the teachers, depending on the grade or subject matter they teach, "maximize classroom time in order to adopt various methods of instruction such as active learning strategies, peer instruction, problem-based learning, or mastery or Socratic methods" (Hamdan et al., 2013, 6).

To conclude, flipping EFL classroom is more than simply replacing traditional face-to-face lectures with narrated PowerPoint videos, instructional videos or ready videos or materials from a wide range of websites. Instead, as a response to the advent of new technologies that pervade all aspects of today's life, it is a call for a shift in the traditional pedagogy by moving large group learning space into the individual learning space independent of time and place thanks to the integration of some of those technologies into learning and teaching procedures.

Regarding all these aforementioned issues about the flipped classroom instruction, the model turns out to be not as a cure for all types of educational problems but a strong potential to promote learning with a careful and an effective planning that will also embrace its possible constraints.

All in all, the present study enlightens the concerned bodies in practice about all aspects of planning and implementation procedures for flipping an EFL classroom. In this respect, the present study serves as a proposal for flipped classroom as a viable pedagogical model that will both respond to learning and teaching needs in the 21st century and boost the psychological predictors of EFL achievement at higher education institutions in Turkey.

5.2.2. Implications for Further Research

The current research has sought to probe the impact of the flipped classroom instruction on EFL performance of university students enrolled at English Prep Schools of universities in Turkish context in addition to its strong affective predictors as attitudes and self-efficacy beliefs. In this respect, it has aimed to contribute to the present literature about flipping EFL classrooms. Within the findings and limitations of this study, this section draws the boundaries of recommendations for further research.

Foremost, it is a must to take a cautionary judgment about generalizing the findings of the study since the results of each study are specific to the academic context in which they are conducted. To the researcher's scope of study, the significant impact of flipped classroom on EFL performance compared to traditional teaching approach was strongly emphasized in the related literature although there seemed no single study to adopt performance in all skill areas in EFL. On the other hand, there seem no consistent efforts to indicate its impact on the retention of that performance. Furthermore, no explicit reference was found to be made to neither its comparative effect on attitudes toward and self-efficacy beliefs in EFL. These blanks in the related literature make the replications of this current study following the similar research design the leading recommendation put forward by the present researcher since such studies may indicate some results different than those found in this study. As a result, those probable differences in the results may help the concerned bodies come up with consistent results to draw general conclusions on the role of flipped classroom in EFL in higher education institutions in Turkey. Such a comprehensive understanding of flipping EFL classroom may help us ground EFL instruction in better empirical research.

Another implication for further researchers lays in the remarks of Hung (2015) who drew attention to surplus of research on the flipped classroom in STEM disciplines and Horn (2013) who emphasized that it was mostly applied in K-L2 education setting. Based on these points, the related literature needs to be expanded to gain insights about the impact of flipped classroom on non-STEM higher education settings (Abeysekera, Dawson, 2015; Hung, 2015). Similarly, these studies have been implemented with university students and these researches in Turkey were generally limited to science education, mathematics education and classroom teacher candidates. Following the assertion made by Bergmann and Sams (2012) that flip classroom model is theoretically claimed to be applicable in any subject area with students from different levels, further researchers are recommended to study the efficiency of flipping EFL classroom at university level by increasing the sample size with students from universities all around the country. This will also eliminate a limitation of the current study regarding the sample size since it included students studying EFL at English Prep Schools of Gebze Technical University in Kocaeli.

The basic implications drawn above based on the limitations of the current study and the blanks in the related literature aimed to strengthen general conclusion on the impact of flipping EFL classrooms on academic performance, attitudes toward and self-efficacy beliefs in EFL. However, to better reinforce understanding about this recent phenomenon in the pedagogy in 21st century, further research may also investigate:

- ✓ the efficiency of flipped classroom in specific sub-skills of EFL,
- ✓ the efficiency of flipped classroom in EFL in a period longer than a semester with a larger sample,
- ✓ the effect of different atmosphere and materials on the efficiency of flipped classroom model.
- ✓ the effects of flipped classroom on different affective traits other than attidutes and self-efficacy that are assigned to be predictors of success or failure in EFL in Turkey,
- ✓ the impact of some other variables on the efficiency of flipped classroom,
- ✓ the effectiveness of flipped classroom at different educational levels.

- ✓ the effectiveness of flipped classroom according to different disciplines,

- ✓ comparative attitudes toward flipping classrooms in different subject areas at different educational levels,

- ✓ the implementation of flipped classroom into classes with unmotivated students,

- ✓ the impact of teachers' computer literacy on the efficiency of flipped classroom model,

- ✓ teachers', parents' and curriculum designers' perceptions of flipped classroom model,

- ✓ student teachers' awareness of flipped classroom model as an innovative instructional approach for change in pedagogy.

To conclude, the present study has yielded valuable results that must be taken into account by a wide range of concerned bodies to tailor EFL practice in Turkey. On the other hand, thanks to its unique aspects, the study also contributed to the present literature about flipping EFL classrooms. Drawing on all the findings reached and comments made as a result of the study, it is believed that the current study will act as a sparkle for the further research to expand the understanding about flipped classroom in order to improve EFL in Turkey.

In this respect, despite conforming to much of the research in STEM context in terms of the positive outcome of flipped classroom, "no conclusive or generalizable findings on flip teaching can be derived from the currently available literature due to insufficient empirical validation across contexts" (Hung, 2015, 93). Hence, further research needs to be conducted to validate and generalize the results of the study in order to set major benefits and drawbacks of this new instructional design for use with different students from different faculties in EFL higher education contexts.

REFERENCES

Abeysekera, Lakmal, Phillip Dawson. 2015. Motivation and Cognitive Load in The Flipped Classroom: Definition, Rationale and A Call for Research. **Higher Education Research & Development.** c. 34. s. 1: 1-14.

Abidin, Mohamad Jafre Zainol, Majid Pour-Mohammadi, Hanan Alzwari. 2012. EFL Students' Attitudes Toward Learning English Language: The Case of Libyan Secondary School Students. **Asian Social Science**. c. 8. s. 2: 119.

Ahmad, Samah Zakareya. 2016. The Flipped Classroom Model to Develop Egyptian EFL Students' Listening Comprehension. **English Language Teaching.** c. 9. s.9: 166-178.

Aiken, Lewis R. 1979. Attitudes Toward Mathematics and Science in Iranian Middle Schools. **School Science and Mathematics.** c. 79. s.3: 229-234

Akbari, Ramin, Kobra Hosseini. 2008. Multiple Intelligences and Language Learning Strategies: Investigating Possible Relations. **System.** c. 36. s. 2: 141-155.

Akdoğan, Süleyman. 2010. Türkiye'de Yabanci Dil Öğretiminde Karşılaşılan Sorunlar ve Bir Çözüm Önerisi Olarak Yabanci Dil Okullarına Yönelik Öğretmen ve Öğretim Elemanlarının Görüşleri. Yüksek Lisans Tezi. Fırat Üniversitesi Sosyal Bilimler Enstitüsü.

Akgün, Ergün, Erdi O. Yılmaz, S. Sadi Seferoğlu. 2011. Vizyon 2023 Strateji Belgesi Ve Fırsatları Artırma Ve Teknolojiyi İyileştirme Hareketi (FATİH) Projesi: Karşılaştırmalı Bir İnceleme. **Akademik Bilişim.** 2-4.

Akgün, Muhterem. 2015. Ters-Düz Sınıfların Öğrencilerin Akademik Başarısı ve Görüşlerine, Etkisi. Yüksek Lisans Tezi. Fırat Üniversitesi Eğitim Bilimleri Enstitüsü

Akıncı, Ahmet, Meltem Kurtoğlu, Süleyman Sadi Seferoğlu. 2012. Bir Teknoloji Politikası Olarak Fatih Projesinin Başarılı Olması İçin Yapılması Gerekenler: Bir Durum Analizi Çalışması. **Akademik Bilişim.**1-10.

Akkoyunlu, Buket. 2001. Öğretmenlerin İnternet Kullanımlan Üzerine Bir Çalışma. **Journal Of Qafqaz University.** s. 8: 57-66.

Akkuş, Ozlem. 2009. Türkiye'de İngilizce Öğrenim ve Öğretiminde Karşılaşılan Sorunlar (Sivas İli Örneği).Yüksek Lisans Tezi. Cumhuriyet Üniversitesi Sosyal Bilimler Enstitüsü.

Aküzel, Gül. 2006. İlköğretim 4-8. Sınıflarda Yabancı Dil Öğretimindeki Başarısızlık Nedenlerinin İncelenmesi (Adana Örneği).Yüksek Lisans Tezi. Çukurova Üniversitesi Sosyal Bilimler Enstitüsü.

Akyel, A. 2003. Yabanci Bir Dil Olarak İngilizce Eğitimi Ve Öğretiminde Yaşanan Sorunlar, Çözümleri Ve Gelişmeler. **Avrupa Birliğine Giriş Sürecinde Türk Eğitim Sisteminde Yabancı Dil Eğitimi Ve Kalite Arayışları.** ed.İrfan Erdoğan. İstanbul: Özel Okullar Derneği Yayınları: 97-102.

Al-Harbi, Sarah S.,Yousif A. Alshumaimeri. 2016. The Flipped Classroom Impact in Grammar Class on EFL Saudi Secondary School Students' Performances and Attitudes. **English Language Teaching.** c. 9. s. 10: 60-80.

Altun, Sadegül Akbaba, Hale Ilgaz. 2016. A Qualitative Case Study of Tablet Use in a High School in Turkey. **ICT in Education in Global Context**, Singapore: Springer: 37-51.

Anderson, Virginia, Barbara Walvoord. 1998. **Effective Grading: A Tool for Learning and Assessment.** San Francisco CA: Jossey-Bass.

Arnold-Garza, Sara. 2014. The Flipped Classroom Teaching Model and Its Use for Information Literacy Instruction. **Communications in Information Literacy.** c. 8. s. 1: 7-22.

Ash, Katie. 2012. Educators View 'Flipped' Model with A More Critical Eye. **Education Week.** c. 32. s. 2: 6-7.

Aydın, Betül. 2016. Ters Yüz Sınıf Modelinin Akademik Başarı, Ödev/Görev Stres Düzeyi ve Öğrenme Transferi Üzerindeki Etkisi. Yüksek Lisans Tezi. Süleyman Demirel Üniversitesi Eğitim Bilimleri Enstitüsü.

Ayhan, Yusuf, Ziya. 1999. Orta ve Yüksek Öğretimde Yabancı Dil Öğretimindeki Başarısızlığın Nedenleri. Yüksek Lisans Tezi. Kocatepe Üniversitesi Sosyal Bilimler Enstitüsü.

Bademci, Vahit. 2011. Kuder-Richardson 20, Cronbach'in Alfası, Hoyt'un Varyans Analizi, Genellenirlik Kuramı ve Ölçüm Güvenirliği Üzerine Bir Çalışma. **Dicle University Journal of Ziya Gokalp Education Faculty**. s. 17: 173-193.

Baker, J. Wesley. 2000. The "Classroom Flip": Using Web Course Management Tools to Become the Guide by the Side. **11th International Conference on College Teaching and Learning**. USA: Jacksonville, Florida.

Baker, William, Kamonpan Boonkit. 2004. Learning Strategies in Reading and Writing: Eap Contexts. **RELC Journal.** c. 35. s. 3: 299-328.

Balıkçı, Hasan Celal. 2015. 'Flipped Classroom' Modeliyle Hazırlanan Derse İlişkin Öğrenci Görüşlerinin ve Ders Başarılarının Değerlendirilmesi. Yüksek Lisans Tezi. Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsü.

Balcı, Ali. 2005. **Sosyal Bilimlerde Araştırma: Yöntem, Teknik ve İlkeler.** Ankara: Pegem Akademi Yayıncılık.

Bandura, Albert. 1978. The Self-System in Reciprocal Determinism. **American Psychologist.** s. 33: 344–358.

_____. 1994. Self-Efficacy. **Encyclopedia of Human Behavior.** 5th ed. S. Ramachaudran. New York: Academic Press. 71-81.

_____. 1997. **Self-Efficacy: The Exercise of Control**. New York: Freeman

Banitt, Justin, Sharon Theis, Lucas Van Leeuwe. 2013. **The Effects Of Technology Integration on Student Engagement.** Minnesota: St. Catherine University.

Basal, Ahmet. 2015. The Implementation of a Flipped Classroom in Foreign Language Teaching. **Turkish Online Journal of Distance Education**. c. 16. s. 4: 28-37.

Başaran, Süleyman, Neşe Cabaroğlu. 2015. The Effect of Language Learning Podcasts on English Self-Efficacy. **International Journal of Language Academy.** c. 2. s. 2: 48-69.

Bektas, H.,&Akman, S. U. (2013). Yükseköğretimde Hizmet Kalitesi Ölçeği: Güvenilirlik Ve Geçerlilik Analizi. Ekonometri Ve İstatistik Dergisi, (18), 116-133.

Bergmann, Jonathan, Aaron Sams. 2012. **Flip Your Classroom: Reach Every Student in Every Class Every Day.** Alexandria, VA: International Society For Technology in Education; ASCD.

Bishop, Jacob Lowell, Matthew A. Verleger. 2013. The Flipped Classroom: A Survey of The Research. **ASEE National Conference Proceedings**. c. 30. s. 9: 1-18.

Bland, J. Martin, Douglas G. Altman. 1997. Statistics Notes: Cronbach's Alpha. **BMJ**. c. 314: 572.

Block, Ellen. 1986. The Comprehension Strategies of Second Language Readers. **TESOL Quarterly**. c. 20. s. 3: 463-492.

Bloom, Benjamin. S. 1995. **İnsan Nitelikleri ve Okulda Öğrenme.** çev. Dursun Ali Özçelik. Istanbul: Turkish National Education Ministry.

Boyraz, Serkan. 2014. İngilizce Öğretiminde Tersine Eğitim Uygulamasının Değerlendirilmesi. Yüksek Lisans Tezi. Afyon Kocatepe ÜniversitesiSosyal Bilimler Enstitüsü.

Brantmeier, Cindy. 2006. Adult Second Language Reading in the USA: The Effects of Readers' Gender and Test Methods. **Forum on Public Policy Online**.c. 2006. s. 1: 1-36.

Brame, Cynthia. J. 2013. Flipping the Classroom. https://s3.amazonaws.com/vu-wp0/wp-content/uploads/sites/59/2017/06/08123735/Flipping-theClassroom1.pdf[14.03.2016].

Bryant, Fred B.,Paul R. Yarnold. 1995. Principal-Components Analysis and Exploratory and Confirmatory Factor Analysis. **Reading and Understanding Multivariate Statistics.** ed, Laurence G. Grimm, Paul R.Yarnold Washington, DC, US: American Psychological Association: 99-136.

Büyüköztürk, Şener. 2007a. **Sosyal Bilimler İçin Veri Analizi El Kitabı: İstatistik, Araştırma Deseni, SPSS Uygulamaları ve Yorum.** Ankara: Pegem Akademi Yayıncılık

_____. 2007b. **Deneysel Desenler: Öntest Sontest Kontrol Gruplu Desen Ve Veri Analizi**. Ankara: Pegem Yayınları,

Cenoz, J. 2001. Three Languages in Contact: Language Attitudes in The Basque Country**. Language Awareness in the Foreign Language Classroom**. ed. Herrarte, David Lasagabaster, Juan Manuel Sierra Zarauz. Zarautz: University of the Basque Country: 37-60.

Chamot, Anna Uhl. 2005. Language Learning Strategy Instruction: Current Issues and Research. **Annual Review of Applied Linguistics**. s. 25: 112-130.

Chan, Jason C.K. 2009. When Does Retrieval Induce Forgetting and When Does It Induce Facilitation? Bergmanımplications for Retrieval Inhibition, Testing Effect, and Text Processing. **Journal of Memory and Language.** c. 61. s. 2: 153-170.

Chao, Chih-Yang, Yuan-Tai Chen, Kuei-Yu Chuang. 2015.Exploring Students' Learning Attitude and Achievement in Flipped Learning Supported Computer Aided Design Curriculum: A Study in High School Engineering Education. **Computer Applications in Engineering Education.** c. 23. s. 4: 514-526.

Chen, So-Chen, Stephen J. H. Yang, Chia-Chang Hsiao. 2015. Exploring Student Perceptions, Learning Outcome And Gender Differences in A Flipped Mathematics Course. **British Journal of Educational Technology**. c. 47. s. 6: 1096-1112.

Child, Dennis. 1990. **The Essentials of Factor Analysis.** Cassell Educational.

Clément, Richard, Zoltán Dörnyei, Kimberly A. Noels. 1994.Motivation, Self-Confidence, and Group Cohesion in The Foreign Language Classroom. **Language Learning.** c. 44. s. 3: 417-448.

Cohen, Jacob. 1988. **Statistical Power Analysis for the Behavioral Sciences.** 2nd ed. Hillsdale, NJ: Erlbaum.

_____. 1994.The Earth Is Round (P < .05). **America Psychologist**. s. 49: 1304-1312.

Cole, Jane E., Jeffrey B. Kritzer. 2009. Strategies For Success: Teaching An Online Course. **Rural Special Education Quarterly.** c. 28 s. 4: 36-40.

Cortina, Jose M. 1993. What Is Coefficient Alpha? An Examination of Theory and Applications. **Journal Of Applied Psychology.** c.78. s. 1: 98-104

Cotterall, Sara. 1999. Key Variables in Language Learning: What Do Learners Believe About Them? **System.** c. 27. s. 4: 493-513.

Creswell, John W., Dana L. Miller. 2000. Determining Validity in Qualitative Inquiry. **Theory And Practice.** c. 39. s. 3: 124-131.

Creswell, John W. 2012. **Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research.** 4th ed. Boston, MA: Pearson Publication.

Cronbach, Lee, J. 1951. Coefficient Alpha and The Internal Structure of Tests. **Psychometrika.** c. 16. s. 3: 297-334.

Crouch, Catherine H.,Eric Mazur. 2001. Peer Instruction: Ten Years of Experience and Results. **American Journal of Physics**. c. 69. s. 9: 970-977.

Crozier, W. Ray. 1997. **Individual Learners: Personality Differences in Education.** London: Routledge.

Crystal, David. 1997. **English As A Global Language.** Cambridge: Cambridge University Press.

Csizér, Kata, Zoltán Dörnyei. 2005. The Internal Structure of Language Learning Motivation and Its Relationship with Language Choice and Learning Effort. **The Modern Language Journal.** c. 89. s. 1: 19-36.

Çakır, Özler, Devrim Alıcı. 2009. Seeing Self As Others See You: Variability in Self-Efficacy Ratings in Student Teaching. **Teachers and Teaching: Theory and Practice.** c. 15. s. 5: 541–561.

Çapık, Canturk. 2014. Geçerlik ve Güvenirlik Çalışmalarında Doğrulayıcı Faktör Analizinin Kullanımı. **Anadolu Hemşirelik ve Sağlık Bilimleri Dergisi.** c.17. s. 3: 196-205.

Çatal, Özlem Gök. 2015. İlköğretim 6, 7 Ve 8 Sınıflarda İngilizce Öğrenme ve Öğretme Sürecinde Karşılaşılan Sorunların İncelenmesi: Öğretmen ve Öğrenci Görüşleri. Yüksek Lisans Tezi. Mehmet Akif Ersoy Üniversitesi Eğitim Bilimleri Enstitüsü.

Çekmez, Erdem, Cemalettin Yıldız, Suphi Önder Bütüner. 2012. Phenomenographic Research Method. **Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education.** c. 6. s. 2: 77-102.

Çelebi, Songul. 2009. Teachers and Students' Views on Anxiety in English Classrooms and Attitudes toward English. Yüksek Lisans Tezi. Çukurova Üniversitesi Sosyal Bilimler Enstitüsü.

Çetinkaya, Yeşim Bektaş. 2009. Language of Others: EFL Students' Perception of and Attitude toward English. **Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi.** c. 18. s. 1: 109-120.

Çimen, Sevda. 2011. Eğitim Fakültesi Öğrencilerinin İngilizceye Yönelik Tutum, İngilizce Kaygısı ve Öz Yeterlik Düzeylerinin İncelenmesi. Yüksek Lisans Tezi. Zonguldak Karaelmas Üniversitesi Sosyal Bilimler Enstitüsü.

Davies, Randy S., Douglas L. Dean, Nick Ball. 2013. Flipping The Classroom and Instructional Technology Integration in a College-Level Information Systems Spreadsheet Course. **Educational Technology Research And Development.** c. 61. s. 4: 563-580.

Day, Jason A.,James D. Foley. 2006. Evaluating A Web Lecture Intervention in A Human–Computer Interaction Course. **IEEE Transactions on Education.** c. 49. s. 4: 420-431.

Demir, Huseyin. 2015. Speaking Anxiety Among Turkish EFL Students (Case of IBSU). **Journal of Education**. c. 4. s.1: 37-43.

Dilbaz, Gülnihal Alkan, Sinan Özgelen, Tuğba Yanpar Yelken. 2012. Araştırma Becerileri Testinin (Abt) Geliştirilmesi. **Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi.** 305-332.

Doordinejad, Farhad Gorban, Hamideh Afshar. 2014. On The Relationship Between Self-Efficacy and English Achievement among Iranian Third Grade High School Students. **International Journal of Language Learning and Applied Linguistics World.** c. 6. s. 4: 461-470.

Douglas, Onika, Kimberly Smith Burton, Nancy Reese-Durham.2008. The Effects of the Multiple Intelligence Teaching Strategy on the Academic Achievement of Eighth Grade Math Students. **Journal of Instructional Psychology.** c. 35. s. 2: 182-187.

Downes, John M.,Penny Bishop. 2012. Educators Engage Digital Natives and Learn from Their Experiences with Technology. **Middle School Journal.** c. 43. s. 5: 6-15.

Dörnyei, Zoltán. 1998. Motivation in Second and Foreign Language Learning. **Language Teaching.** c. 31. s. 3: 117-135.

Dörnyei, Zoltan, István Ottó. 1998. Motivation in Action: A Process Model of L2 Motivation. **Working Papers in Applied Linguistics.** s. 4: 43-69.

Duy, Le Hai. 2016. Improving EFL Achievement Tests at High Schools. **International Journal of Humanities & Social Science Studies.** c. 2. s. 5: 143-156

Educause L. I. 2012.7 Things You Should Know about Flipped Classrooms. EDUCAUSE Creative Commons. https: //net.educause. Edu /ir/ library /pdf /eli 7081.pdf[16.04.2015].

Ekmekçi, Emrah. 2014. Flipped Writing Class Model with a Focus on Blended Learning. Phd Thesis. Gazi University Educational Sciences Institute.

Engin, Ali Osman, Mehmet Ali Seven. 2007. Factors Which Affect the Success in English Teaching in Turkey (The Example Of Erzurum). Online Submission.http://files.eric.ed.gov/fulltext/ED497452.pdf [16.04.2015].

Engin, Marion. 2014. Extending The Flipped Classroom Model: Developing Second Language Writing Skills Through Student-Created Digital Videos. **Journal of the Scholarship of Teaching and Learning.** c.14. s. 5: 12-26.

Erişen, Yavuz. 2007. Determination of the Quality Assurance Standards in Terms Of EU Priorities, For Teacher Training Programs in Turkey. **9th International Conference European Unification and Educational Challenges in The Balkans.** Greece:

Erkan, Dilek Yavuz.,& Saban, Ayten I. 2011. Writing Performance Relative to Writing Apprehension, Self-Efficacy in Writing, and Attitudes toward Writing: A Correlational Study in Turkish Tertiary-Level EFL. **The Asian EFL Journal Quarterly.** c. 13. s. 1: 163-191.

Ertürk, Selahattin. 1972. **Eğitimde Program Geliştirme.** Ankara: Hacettepe Üniversitesi Basımevi.

Excellence C. F. 2013. Flipping Your Course. https://Dakinburdick.Files.Wordpress. Com/2013/09/Flippingyourcoursev2.Pdf[17.06.2015].

Fakeye, David O. 2010. Students' Personal Variables as Correlates of Academic Achievement in English As ASecond Language in Nigeria. **Journal of Social Sciences.** c. 22. s. 3: 205-211.

Farah, Mireille. 2014. The Impact of Using Flipped Classroom Instruction on The Writing Performance of Twelfth Grade Female Emirati Students in The Applied Technology High School (ATHS). Phd Thesis. The British University

Fives, Helenrose, Nicole DiDonato-Barnes. 2013. Classroom Test Construction: The Power of A Table of Specifications. **Practical Assessment, Research & Evaluation.** c.18 s. 3: 1-7.

Fulton, Kathleen P. 2013. Byron's flipped classrooms. **The Education Digest**. c. 79. s. 1: 22-26.

Gallagher, Kathleen. 2009. From Guest Lecturer To Assignment Consultant: Exploring A New Role For The Teaching Librarian. **LOEX Conference Proceedings 2007:** 39-43.

Gannod, Gerald C., Janet E. Burge, Michael T. Helmick. 2008. Using the inverted classroom to teach software engineering**. Proceedings of The 30th International Conference on Software Engineering**, **May 10-18 2008. Leipzig, Germany:** 777-786.

Gardner, Howard. 1997. Multiple Intelligences as a Partner in School Improvement. **Educational Leadership.** c. 55. s. 1: 20-21

Gardner, Robert C., Wallace E. Lambert. 1972. **Attitudes and Motivation in Second Language Learning**. USA: Newbury House

Gardner, Robert C. 1985. **Social Psychology and Second Language Learning: The Role of Attitudes and Motivation.** London: Edward Arnold.

Gardner, Robert C. 2006. The Socio-Educational Model of Second Language Acquisition: A Research Paradigm. **Eurosla Yearbook.** c. 6. s. 1: 237-260.

Gastwirth, Joseph, Yulia R. Gel, Weiven Miao. 2009. The impact of Levene's test of equality of variances on statistical theory and practice. **Statistical Science**. c. 24. s. 6: 343-360.

Gençer, Büşra Gamze. 2015. Okullarda Ters-Yüz Sınıf Modelinin Uygulanmasına Yönelik Bir Vaka Çalışması. Yüksek Lisans Tezi. Bahçeşehir Üniversitesi Eğitim Bilimleri Enstitüsü.

Genç, Gülten, Emine Kuluşaklı, Savaş Aydın. 2016. Exploring EFL Learners' Perceived Self-Efficacy and Beliefs on English Language Learning. **Australian Journal Of Teacher Education.** c. 41. s. 2: 52-68.

Ghasemboland, Farimah, Fatimah Binti Hashim. 2013. Teachers' Self-Efficacy Beliefs and Their English Language Proficiency: A Study of Nonnative EFL Teachers in Selected Language Centers. **Procedia-Social and Behavioral Sciences**. c. 103: 890-899.

Gilakjani, Abbas Pourhosein, Narjes Banou Sabouri, Asieh Zabihniaemran. 2015. What Are The Barriers in The Use Of Computer Technology in EFL Instruction? **Review Of European Studies.** c. 7. s. 11: 213-221.

Gilboy, Mary Beth, Scott Heinerichs, Gina Pazzaglia. 2015. Enhancing Student Engagement Using the Flipped Classroom. **Journal of Nutrition Education and Behavior.** c. 47. s. 1: 109-114.

Golafshani, Nahid. 2003. Understanding Reliability and Validity in Qualitative Research. **The Qualitative Report.** c. 8. s. 4: 597-606.

Gömleksiz, Mehmet Nuri. 1993. Yüksek Öğretimde Yabancı Dil Öğretimi ve Sorunları (Fırat Üniversitesi Örneği). Yüksek Lisans Tezi. Fırat Üniversitesi Sosyal Bilimler Enstitüsü

_____. 2010. An Evaluation of Students' Attitudes Toward English Language Learning in Terms of Several Variables. **Procedia Social And Behavioral Sciences.** c. 9. s. 1: 913-918

Graham, Charles. R. 2006. Blended Learning Systems.**The Handbook Of Blended Learning: Global Perspectives, Local Designs.**ed. Bonk, Curtis J.,Charles R. Graham. USA: Pfeiffer: 3-21.

Gorsuch, Greta. 2009. Investigating Second Language Learner Self-Efficacy and Future Expectancy of Second Language Use for High-Stakes Program Evaluation. **Foreign Language Annals**. c.42. s. 3: 505-540.

Gravetter, Frederick J., Larry B. Wallnau. 2014. **Essentials of Statistics for the Behavioral Sciences.** 8th ed. Belmont, CA: Wadsworth.

Guba, Egon G., Yvonna S. Lincoln. 1982. Epistemological and Methodological Bases of Naturalistic Inquiry. **Educational Communication and Technology Journal.** c. 30. s. 4: 233-252.

Guilherme, Manuela. 2007. English As A Global Language and Education for Cosmopolitan Citizenship. **Language and Intercultural Communication.** c.7. s. 1: 72-90.

Gülmez, Yener. 1982. Yabancı Dil Öğretiminde Öğrenci Başarısını Etkileyen Faktörler. Doktora Tezi. Ankara: Ankara Üniversitesi Eğitim Fakültesi.

Gürer, Melih Derya, Erkan Tekinarslan, Nuh Yavuzalp. 2016. Opinions of Instructors Who Give Lectures Online about Distance Education. **Turkish Online Journal of Qualitative Inquiry.** c. 7. s. 1: 47-78.

Haley, Marjorie Hall. 2004. Understanding Learner-Centered Instruction from Perspective Of The Multiple Intelligences. **Foreign Language Annals.** c. 34. s. 4: 355-367.

Hamdan, Noora.,Patrick Mcknight, Katherine Mcknight, & Kari M. Arfstrom. 2013. A Review of Flipped Learning. Pearson-Research & Innovation Network http://ResearchNetwork.Pearson.Com/Wpcontent/Uploads/Litreview_Flippedlearning1pdf [11.11.2015].

Han, Yu Jung. 2015. Successfully Flipping the ESL Classroom For Learner Autonomy. **NYS TESOL Journal.** c. 2. s. 98: 98-109.

Hanover Research. 2013. Best Practices for The Flipped Classroom. K-12 Blog: http://www.Hanoverresearch.Com/2013/10/Best-Practices-For-The-Flipped Classroom [11.06.2016].

Hao, Yungwei. 2016. Middle School Students' Flipped Learning Readiness in Foreign Language Classrooms: Exploring Its Relationship with Personal Characteristics And Individual Circumstances. **Computers in Human Behavior.** c. 59: 295-303.

Harrington, Donna. 2008. **Confirmatory Factor Analysis.** USA: Oxford University Press.

Harvill, Leo. M. 1991. Standard Error of Measurement. **Educational Measurement: Issues and Practice.** c. 10. s. 2: 33-41.

Headden, Susan. 2013. The Promise of Personalized Learning: Blending the Human Touch with Technological Firepower. **Education Next.** c. 13. s. 4: 14-21.

Horn, Michael B. 2013. The Transformational Potential of Flipped Classrooms. **Education Next.** c.13. s. 3: 78-79.

Hosenfeld, Carol. 1977. A Preliminary Investigation of the Reading Strategies of Successful and Nonsuccessful Second Language Learners. **System.** c. 5. s. 2:110-123.

Herold, Benjamin. 2016.Technology in Education: An Overview. **Education Week**. http://Www.Edweek.Org/Ew/Issues/Technology-in-Education[08.05.2016].

Hung, Hsiu-Ting. 2015. Flipping the Classroom For English Language Learners To Foster Active Learning. **Computer Assisted Language Learning.** c. 28. s. 1: 81-96.

Hycner, Richard H. 1985. Some Guidelines for the Phenomenological Analysis of Interview Data. **Human Studies.** c.8. s. 3: 279-303.

Ibrahim, Mohamed, Rebecca Callaway. 2014. Students' Learning Outcomes and Self-Efficacy Perception in A Flipped Classroom. **Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education**. Association for the Advancement of Computing in Education (AACE):14-23.

İlhan, Dilek, Hakan Karataş. 2015. An Analysis on Motivational Beliefs and Attitudes of Undergraduates Regarding Learning English. **International Journal of Educational Researcher.** c. 6. s. 2: 1-14.

İnceçay, Görsev, Ayşe S. Akyel. 2014. Turkish EFL Teachers' Perceptions of English As A Lingua Franca. **Turkish Online Journal of Qualitative Inquiry.** c. 5. s. 1: 1-19.

Internet World Stats. [08.03.2016].Internet World Users by Language. http://Www.internetworldstats.Com/Stats7.Htm

Ivanova, Vanya. 2011. Construction and Evaluation of Achievement Tests in English. **National Conference, Education in The Information Society:** 276-285.

İyitoglu, Orhan, Hasan Aydin. 2015. The Relationship between Multiple Intelligence Profiles and Reading Strategy Use of Successful English as A Foreign Language (EFL) Readers. **SA Journal of Education.** c. 35. s. 2: 1-11.

Javier, Vivaldo-Lima. 1997. Cognitive Style And Reading Comprehension in L1 and L2: **The 31st Annual TESOL Convention.** Orlando, Florida: 1-9.

Jenkins, Jennifer. 2008. EFL at the Gate: The Position of English as a Lingua Franca Ideas from the Corpora. **Humanising Language Teaching.** c. 17. s. 2.

Johnson, Graham Brent. 2013. Student Perceptions of the Flipped Classroom. Doctoral Dissertations. The University of British Columbia Graduate School of Educational Technology.

Jones, Edward E. 1979. The Rocky Road From Acts to Dispositions. **American Psychologist.** c. 34. s. 2: 107-117.

Kabaharnup, Çiğdem. 2010. The Evaluation of Foreign Language Teaching in Turkey: English Language Teachers' Point of View. Master Thesis. Çukurova University Institute of Social Sciences.

Kaiser, Henry F. 1960. The Application of Electronic Computers to Factors Analysis. **Educational and Psychological Measurement.** c. 20. s. 1: 141-151.

Kajornboon, Annabel Bhamani. 2005. Using Interviews as Research Instruments**. E-Journal for Research Teachers.** c. 2. s. 1: 1-9.

Kanadlı, Sedat, Birsen Bağçeci. 2015. Investigation of Students' Self Efficacy Beliefs in Terms of Perceived Autonomy Support. **Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi.** c. 12. s. 30: 98-112.

Kang, Namhee. 2015. The Comparison between Regular and Flipped Classrooms for EFL Korean Adult Learners. Multimedia- Assisted Language Learning, c. 18. s. 3: 41-72.

Kara, Ahmet. 2009. The Effect of A 'Learning Theories' Unit on Students' Attitudes Toward Learning. **Australian Journal of Teacher Education**. c. 34. s. 3: 100-113.

Kara, Cüneyt Orhan. 2016a.Tıp Fakültesi Klinik Eğitiminde Ters Yüz Sınıf Modeli Kullanılabilir Mi? Yüksek Lisans Tezi. Akdeniz Üniversitesi, Sağlık Bilimleri Enstitüsü.

_____. 2016b. Ters Yüz Sınıf. **Tıp Eğitimi Dünyasi.** c. 15. s. 45: 12-26.

Karahan, Firdevs. 2007. Language Attitudes of Turkish Students toward The English Language and Its Use in Turkish Context. **Journal of Arts and Sciences.** s. 7: 73-87.

Karakoç, Fatma Yeşim, Levent Dönmez. 2014. Ölçek Geliştirme Çalışmalarında Temel İlkeler. **Tıp Eğitimi Dünyası.** c. 40. s. 40: 39-49.

Karasar, Niyazi. 2007. **Bilimsel Araştırma Yöntemi**.17. bs. Ankara: Nobel Yayınevi.

Kazazoğlu, Selim. 2013. Türkçe ve İngilizce Derslerine Yönelik Tutumun Akademik Başarıya Etkisi. **Eğitim Ve Bilim. c.** 38. s. 170: 294-307.

Keshavarz, Mohammad Hossein, Soroor Ashtarian. 2008. The Relationship Between Iranian Efl Learners' Gender And Reading Comprehension of Three Different Types of Text. **IJAL.** c. 11. s. 1: 97-113

Khajavi, Yaser, Saeed Ketabi. 2012. Influencing EFL Learners' Reading Comprehension and Self-Efficacy Beliefs: The Effect of Concept Mapping Strategy**. Porta Linguarum: Revista İnternacional De Didáctica De Las Lenguas Extranjeras.** s. 17: 9-27.

Khajavy, Gholam Hassan, Behzad Ghonsooly, Azar Hosseini Fatemi. 2017. Testing A Burnout Model Based on Affective-Motivational Factors among EFL Teachers. **Current Psychology.** c. 36. s. 2: 339-349.

Khan Academy. 2012. https://www.Khanacademy.Org/

Kim, MinKyu, So Mi Kim, Otto Khera, Joan Getman. 2014. The Experience of Three Flipped Classrooms in an Urban University: An Exploration of Design Principles. **The Internet and Higher Education.** s. 22: 37-50.

Kim, Do-Hong, Chuang Wang, Hyun Seon Ahn, Mimi Bong. 2015. English Language Learners' Self-Efficacy Profiles and Relationship with Self-Regulated Learning Strategies. **Learning and Individual Differences.** c. 38: 136-142.

King, Allison. 1993. From Sage on The Stage to Guide on The Side. **College Teaching.** c. 41. s. 1:30-35.

Kitchakarn, Orachorn. 2015. EFL Learners' Attitudes toward Using Computers As A Learning Tool in Language Learning. **TOJET: The Turkish Online Journal of Educational Technology.** c. 14. s. 2: 52-58.

Kiziltepe, Zeynep. 2000. Attitudes and Motivation of Turkish EFL Students toward Second Language Learning. ITL. **Institute Voor Togepaste Linguistik.** c. 129. s. 30: 141-168.

Koo, Cathy. L.,Demps, Elaine L., Charlotte Farris, John D. Bowman, J. D., Ladan Panahi, Paul Boyle. 2016. Impact of Flipped Classroom Design on Student Performance and Perceptions in a Pharmacotherapy Course. **American Journal of Pharmaceutical Education.** c. 80. s. 2: 1-9.

Kostka, Ilka, R. Brinks Lockwood. 2015. What's on The Internet for Flipping English Language Instruction? **Tesl-Ej.** c. 19. s. 2: 1-12.

Kul, Sevsl. 2014. Guideline for Suitable Statistical Test Selection. **Plevra Bülteni.** c. 8. s. 2: 26-29.

Kurt, Gökçe. 2017. Implementing the Flipped Classroom in Teacher Education: Evidence From Turkey. **Educational Technology & Society.** c. 20. s. 1: 211-221.

Kvashnina, Olga Sergeevna, Ekaterina Andreevna Martynko. 2016. Analyzing The Potential of Flipped Classroom in ESL Teaching. **International Journal of Emerging Technologies in Learning.** c. 11. s. 3: 71-73.

Lage, Maureen J., Glenn J. Platt, Michael Treglia. 2000. Inverting the Classroom: A Gateway to Creating An Inclusive Learning Environment. **The Journal of Economic Education.** c. 31. s. 1: 30–43.

Lasagabaster, David, Juan Manuel Sierra. 2009. Language Attitudes in CLIL and Traditional EFL Classes. **International CLIL Research Journal.** c. 1. s. 2: 4-17.

Latifah, A. Latif., Mansor Fadzil, Ramli Bahroom., Wardah Mohamad, & Ng ManSan. 2011. The Role of Motivation, Attitude, Anxiety, and Instrumental Orientation in Influencing Learners' Performance in English as Second Language in OUM. **Global Learn**. c. 1: 1659-1668.

Leis, Adrian, Simon Cooke, Akihiko Tohei.2015. The Effects of Flipped Classrooms on English Composition Writing in An EFL Environment. **International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT).** c. 5. s. 4: 37-51

Li, Charlene, Josh Bernof, Cynthia Pflaum, Sarah Glass 2007. How Consumers Use Social Networks. **For Interactive Marketing Professionals.** c. 21: 1-11.

Liliequist, Malin. 2013. Motivation among English L2 Learners in Sweden: A Comparison Between High and Low Achievers in Upper Secondary School. Project in Linguistics. Karlstads Universtiy Faculty of Arts and Education.

Lilliefors, Hubert W. 1967. On The Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. **Journal of The American Statistical Association.** c. 62. s. 318: 399-402.

Liu, Meihua. 2007. Chinese Students' Motivation to Learn English At The Tertiary Level. **Asian EFL Journal.** c. 9. s. 1: 126-146.

Locke, Edwin A., and Gary P. Latham.2006. New Directions in Goal-Setting Theory. **Current Directions in Psychological Science.** c. 15. s. 5: 265-268.

Long, Taotao, Joanne Logan, John Cummins, Michael Waugh. 2016. Students' and Instructor's Attitudes and Receptions of The Viability of Using A Flipped Classroom Instructional Model in A Technology-Enabled Active Learning (TEAL) Classroom. **Journal of Teaching and Learning with Technology.** c. 5.s. 1: 46-58.

Love, Betty, Angie Hodge, Neal Grandgenett, Andrew W. Swift. 2014. Student Learning and Perceptions in A Flipped Linear Algebra Course. **International Journal of Mathematical Education in Science And Technology.** c. 45. s. 3: 317-324.

Lyon, Aidan. 2014. Why Are Normal Distributions Normal? **The British Journal for The Philosophy of Science.** c. 65. s. 3: 621-649.

Marasinghe, Mervyn G., William Q. Meeker, Dianne Cook, Tae-Sung Shin. 1996. Using Graphics and Simulation to Teach Statistical Concepts. **The American Statistician**. c. 50. s. 4: 342-351.

Marefat, H. 2003. The Impact of Teaching Direct Learning Strategies on the Retention of Vocabulary by EFL Learners. **The Reading Matrix.** c. 3. s. 2: 47-62.

Martínez, Uribe Diego, José Gutiérrez Pérez, Daniel Madrid. 2013. Attitudes of Mexican American Students Toward Learning English As A Second Language

in a Structured Immersion Program. **Porta Linguarum: Revista Internacional De Didáctica De Las Lenguas Extranjeras.** s. 20: 205-221

Mazur, Eric. 1997. Peer Instruction: A User's Manuel. New Jersey: Prentice Hall

_____. 2009. Farewell, Lecture. **Science.** c. 323. s. 5910: 50-51.

McGahee, Thayer W., Julia Ball. 2009. How to Read and Really Use an Item Analysis. **Nurse Educator.** c. 34. s. 4: 166-171.

Mcgowan, HerleM. 2011. Planning A Comparative Experiment in Educational Settings. **Journal of Statistics Education.** c. 19. s. 2: 1-18.

Mckeown, John. 2016. From "Sage on The Stage to Guide on The Side": A Case Study of A Transition to Flipped English Language Learning in A Higher Education Setting in Turkey. **ELT Research Journal.** c. 5. s. 2: 145-154.

Mclaughlin, Jacqueline E., La Toya M. Griffin, Denise A. Esserman, Christopher A. Davidson, Dylan M. Glatt, Mary T. Roth, Nastaran Gharkholonarehe, and Russell J. Mumper (2013). Pharmacy Student Engagement, Performance, and Perception in A Flipped Satellite Classroom. **American Journal of Pharmaceutical Education.** c. 77. s. 9: 1-8.

Meriç, Ali. 2015. Foreign Language Teaching Anxiety and Self-Efficacy Beliefs of Turkish Pre-Service EFL Teachers. **The International Journal of Research in Teacher Education.** c. 6. s. 3: 40-58.

Miles, Matthew B., A. Michael Huberman. 1994. **Qualitative Data Analysis: An Expanded Sourcebook.** Thousand Oaks: Sage.

Milman, Natalie B. 2014. The Flipped Classroom Strategy: What is It and How Can It Best Be Used?. **Distance Learning.** c. 11. s. 4: 9-11.

Miller, Andrew. 2012. Five Best Practices for The Flipped Classroom. Edutopia: Http://Www.Edutopia.Org/Blog/Flipped-Classroom-Best-Practices-Andrew-Miller [08.05.2016].

Moghari, Elahe Hejazi,Mas'oud GLavasani, M. G., Valliollah Bagherian, JavadAfshari. 2011. Relationship between Perceivedteacher's Academic Optimism And English Achievement: Role of Self-Efficacy. **Procedia Social and Behavioral Sciences.** c. 15: 2329–2333.

Momani, Mohammed. 2009. The Jordanian Secondary Stage Students Achievement in Reading Comprehension According to Their Views toward Learning English As A Foreign Language. **Journal of Human Sciences.** c. 42. s. 7: 1-36.

Montano, Daniel E., Danuta Kasprzyk. 2008. Theory of reasoned action, theory of planned behavior, and the integrated behavioral model. **Health Behavior and Health Education** Eds. Glanz, K., Rimer, B.K. and Viswanath, K., San Francisco: Jossey-Bass: 67-96

Moravec, Marin, Adrienne Williams, Nancy Aguilar-Roca, Diane K. O'Dowd. 2010. Learn Before Lecture: A Strategy That İmproves Learning Outcomes in A Large Introductory Biology Class. **CBE-Life Sciences Education.** c. 9. s. 4: 473–481

Network, F. L. 2014. The Four Pillars of FLIP™. www.Flippedlearning.Org/ Definition.

Nielsen, Lisa. 2012. Five Reasons I'm Not Flipping Over The Flipped Classroom. **Technology & Learning.** c. 32. s. 10: 46-46.

Novak, Gregor M. 2011. Just-in-Time Teaching. **New Directions for Teaching and Learning.** s. 128: 63-73.

November, Alan, Brian Mull. [14.05.2015]. Flipped Learning: A Response to Five Common Criticisms. http://web.uvic.ca/~gtreloar/Articles/Technology/flipped-learning a-response-to-five-common-criticisms.pdf.

Obari, Hiroyuki, Stephen Lambacher. Successful EFL Teaching Using Mobile Technologies in A Flipped Classroom. **Critical CALL–Proceedings of The 2015 EUROCALL Conference**, 2015: Padova, Italy: 433-438.

Obeidat, Mohammed Mohideen.2005. Attitudes and Motivation in Second Language Learning. **Journal of Faculty of Education.** c. 18. s. 22: 1-17.

Oğuz, Efsun. 1999. İlköğretimde Yabancı Dil (İngilizce) Öğretimi Sorunları. Yüksek Lisans Tezi. Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü.

Osman, Mohamed El Tahir, Hania Al Khamisi, Thuwayba Al Barwani, Abdo Al Mekhlafi. 2016. EFL Reading Achievement: Impact of Gender and Self-Efficacy Beliefs. **International Journal of Learning, Teaching and Educational Research.** c. 15. s. 3: 54-73.

Otrar, Mustafa, Ferhat Süleyman Argın. 2014. Öğrencilerin Sosyal Medyaya İlişkin Tutumlarının Kullanım Alışkanlıkları Bağlamında İncelenmesi. **Journal of Research in Education and Teaching.** c. 3. s. 3: 1-13.

Overmyer, Jerry. 2012. Flipped Classrooms 101. **Principal.** c. 92. s. 1: 46-47.

Oxford, Rebecca L.,Judith A. Burry-Stock.1995. Assessing the Use of Language Learning Strategies Worldwide with The ESL/EFL Version of The Strategy Inventory for Language Learning (SILL). **System.** c. 23. s. 1: 1-23.

Özen, Ahmet. 1979. Yabancı Dil Öğreniminin Gerisindeki Sorunlar. **İzlem Dergisi.** s. 1: 22-25

Park, Hung. Myoung. (2009). **Comparing Group Means: T-tests and One-way ANOVA Using Stata, SAS, R, and SPSS**. Trustees of Indiana University

Patton, Michael Quinn. 1990. **Qualitative Evaluation Methods.** 2nd ed. Thousand Oaks, CA: Sage.

Pedroni, Michela, Bertrand Meyer. 2006. The Inverted Curriculum in Practice. **ACM SIGCSE Bulletin. c.** 38. s. 1: 481-485.

Planning Commission. 1973. Third Five-Year Development Plans, 1973-1977. Ankara: TC Başbakanlık Planning Comission.

Planning Commission. 1973. Fourth Five-Year Development Plan, 1979-1983. Ankara: TC Başbakanlık Planning Comission.

Prince, Michael. 2004. Does Active Learning Work? A Review of The Research. **Journal of Engineering Education.** c. 93. s. 3: 223-231.

Rahimi, Ali, Atiyeh Abedini. 2009. The Interface between EFL Learners' Self-Efficacy Concerning Listening Comprehension and Listening Proficiency. **Novitas-Royal.** c.3. s. 1: 14-28.

Rajagopalan, Kanavillil. 2004. The Concept of 'World English' and Its Implication for ELT. **ELT Journal**. c. 58. s. 2: 111-117.

Raoofi, Saeid, Bee Hoon Tan, Swee Heng Chan. 2013. Self-Efficacy in Second Language Learning Contexts. **English Language Learning.** c. 5. s. 11: 60- 73

Reid, Scott A. 2016. A Flipped Classroom Redesign in General Chemistry. **Chemistry Education Research and Practice.** c. 17. s. 4: 914-922.

Research Trends. 2012. The Language of (Future) Scientific Communication. **Research trends.** c.31. https://www.researchtrends.com/issue-31-november-2012/the-language-of-future-scientific-communication/ [11.10.2016].

Richards, Jack C.,Richard W. Schmidt. 1992. **Longman Dictionary of Language Teaching and Applied Linguistics.** Harlow, UK: Longman.

Rideout, Victoria J., Ulla G. Foehr, Donald F. Roberts. 2010. **Generation M: Media in The Lives of 8-to18-Year-Olds.** California: Henry J. Kaiser Family Foundation.http://files.eric.ed.gov/fulltext/ED527859.pdf [14.10.2015].

Roth, Channy, Suksan Suppasetseree. 2016. Flipped Classroom: Can It Enhance English Listening Comprehension for Pre-Universıty Students in Cambodia? **Proceedings of Clasic: Learning in and beyond the Classroom: Ubiquity in Foreign Language Education**https://www.fas.nus.edu.sg/cls/ CLaSIC/clasic 2016/PROCEEDINGS/roth_channy.pdf255-263 [18.01.2015].

Rukh, Samar. 2014. Students' Attitudes Toward English Language Learning and Academic Achievement: A Case of Business Students in Punjab. **European Academic Research**. c. 2. s. 4: 5596-5612.

Ruscio, John. 2008. A Probability-Based Measure of Effect Size: Robustness to Base Rates and Other Factors. **Psychological Methods**. c. 13. s. 1: 19-30.

Ryan, Barry James. 2013. Flipping Over: Student-Centered Learning and Assessment. **Journal of Perspectives in Applied Academic Practice.** c. 1. s. 2: 30-39.

Sabri, Shafizan. 2013. Item Analysis of Student Comprehensive Test For Research in Teaching Beginner String Ensemble Using Model Based Teaching Among Music Students in Public Universities. **International Journal of Education and Research.** c. 1. s. 12: 1-14.

Sarıtaş, M. Tuncay, Nalan Üner. 2013. Eğitimde Yenilikçi Teknolojiler: Bulut Teknolojisi. **Eğitim ve Öğretim Araştırmaları Dergisi.** c. 2. s. 3: 192-201.

Schumacker, Randall E., Richard G. Lomax. 2004. **A Beginner's Guide to Structural Equation Modeling.** 2nd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Schwarz, Norbert, Gerd Bohner. 2001. The Construction of Attitudes. **Blackwell Handbook of Social Psychology: Intrapersonal Processes.** Malden, Massachusetts: Blackwell Publishers: 436-458.

Shaffer, Shelly. 2016. One High School English Teacher: On His Way to A Flipped Classroom. **Journal of Adolescent & Adult Literacy.** c. 59. s. 5: 563-573.

Sırakaya, Didem Alsancak. 2015.Tersyüz Sınıf Modelinin Akademik Başarı, Öz-Yönetimli Öğrenme Hazırbulunuşluğu ve Motivasyon Üzerine Etkisi. Doktora Tezi. Gazi Üniversitesi Eğitim Bilimleri Enstitüsü.

Siegle, Del. 2014. Technology: Differentiating Instruction by Flipping The Classroom. **Gifted Child Today.** c. 37. s. 1:51-54.

Simon, Marilyn K. 2011.**Dissertation And Scholarly Research: Recipes for Success.** Seattle, WA: Dissertation Success, LLC.

Smith, Larry E. 2015. English as An International Language: No Room For Linguistic Chauvinism. **Journal of English as A Lingua Franca.** c. 4. s. 1: 165-171.

Staker, Heather, Michael B. Horn. 2012. **Classifying K-12 Blended Learning.** Innosight Institute. http://files.eric.ed.gov/fulltext/ED535180.pdf [14.03.2015].

Strayer, Jeremy, F. 2007. The Effects of The Classroom Flip On The Learning Environment: A Comparison of Learning Activity in A Traditional Classroom and A Flip Classroom That Used An Intelligent Tutoring System. Doctoral Dissertation. The Ohio State University

_____ 2009. **Inverting the Classroom: A Study of the Learning Environment When An Intelligent Tutoring System Is Used to Help Students Learn.** Saarbrücken, Germany: VDM-Verlag Müller.

_____ 2012. How Learning in An Inverted Classroom Influences Cooperation, Innovation And Task Orientation. **Learning Environments Research**. c. 15. s. 2: 171-193.

Student. (1908). The Probable Error of s Mean. **Biometrika**. c. 6. s. 1: 1-25.

Şahin, Yusuf. 2009. Yabanci Dil Öğretiminde Öğrenci Başarisini Olumsuz Yönde Etkileyen Unsurlar. **Türkiye Sosyal Araştırmalar Dergisi.** c. 13. s. 1: 149-158

Şentürk, Burcu. 2015. EFL Turkish University Students' Attitudes and Motivation toward Reading in English. **Procedia-Social and Behavioral Sciences.** c. 199: 704-712.

Şimşek, Ömer Faruk. 2007. **Yapısal Eşitlik Modellemesine Giriş: Temel İlkeler ve LISREL Uygulamaları.** Ankara: Ekinoks Yayıncılık

Tavil, Z. Müge 2009. Parental Attitudes Toward English Education for Kindergarten Students in Turkey. **Kastamonu Education Journal.** c. 17. s. 1: 331-34.

Tan, Seref. 2009. Misuses of KR-20 and Cronbach's Alpha Reliability Coefficients. **Egitim ve Bilim.** c. 34. s. 152: 101-111.

Tekin, Halil. 1996. **Eğitimde Ölçme ve Değerlendirme.** 9. bs. Ankara: Yargı Yayınları.

Tenneson, Michael, Bob McGlasson.2006. The Classroom Flip. Missouri Teaching and Learning Mentor Program www.Fontbonne.Edu/Upload/Theclassroom flip [11.11.2016].

Thabane Lehana, Jinhui Ma, Rong Chu, Ji Cheng, Afisi Ismaila, Lorena P. Rios, Reid Robson, Marroon Thabane, Lora Giangregorio, Charles H. Goldsmith. 2010. A Tutorial on Pilot Studies: The What, Why and How. **BMC Medical Research Methodology.** c.10. s. 1: 1-10.

Thompson, Nathan A., Jack Guyer. 2010. **User's Manual for Iteman 4.1.** St. Paul MN: Assessment Systems Corporation.

Tilfarlioğlu, Filiz Talçın, Emrah Cinkara. 2009. Self-Efficacy in EFL: Differences Among Proficiency Groups and Relationship with Success. **Novitas-Royal.** c. 3. s. 2: 129-142.

Todaka, Yuichi. 2013. Self-Efficacy Theory and Beyond: Japanese College EFL Learners. **Journal of Educational and Social Research**. c. 3. s. 7: 359-369.

Toruk, İbrahim. 2008. Üniversite Gençliğinin Medya Kullanma Alışkanlıkları Üzerine Bir Analiz. **Selcuk University Social Sciences Institute Journal**. c. 19: 475-488.

Tsiplakides, Iakovos, Areti Keramida. 2010. Promoting Positive Attitudes in ESL/EFL Classes. **The Internet TESL Journal.** c. 16. s. 1.

Tuncer, Murat, H. Güner Berkant, Yunus Doğan. 2015. İngilizce Dersine Yönelik Tutum Ölçeğinin Geçerlik ve Güvenirlik Çalışması. **Eğitim ve Öğretim Araştırmaları Dergisi.** c. 4. s. 2: 260-266.

Tuncer, Murat, Yunus Doğan. 2016. Relationships Among Foreign Language Anxiety, Academic Self-Efficacy Beliefs and Metacognitive Awareness: A

Structural Equation Modelling. **International Journal of Learning and Development.** c. 6. s. 2: 31-41.

Tunç, Sabiha Özgür. 2003. Use of Language Learning Strategies in Relation to Student Characteristics at Başkent University. Master Thesis. Middle East Technical University Institute of Social Sciences.

Turan, Zeynep. 2015. Ters Yüz Sınıf Yönteminin Değerlendirilmesi ve Akademik Başarı, Bilişsel Yük ve Motivasyona Etkisinin İncelenmesi. Doktora Tezi. Atatürk Üniversitesi Eğitim Bilimleri Enstitüsü

Turgut, Mehmet Fuat, Yaşar Baykul. 2012. **Eğitimde Ölçme ve Değerlendirme**. Ankara: Pegem Akademi.

Turner, J.Rick. & Julian F. Thayer. 2001. **Introduction to analysis of variance.** Thousand Oak, CA: Sage Publications.

Uzunboylu, Hüseyin, Damla Karagozlu. 2015. Flipped Classroom: A Review of Recent Literature. **World Journal on Educational Technology.** c. 7. s. 2: 142-147

Valenza, Joyce Kasman. 2012. The Flipping Librarian. **Teacher Librarian**. c. 40. s. 2: 22-25.

Van Teijlingen, Edwin, Vanora Hundley. 2002. The Importance of Pilot Studies. **Nursing Standard.** c. 16.s. 40: 33-36.

Visser, Max. 2008. Learning Under Conditions of Hierarchy and Discipline: The Case of The German Army, 1939–1940. **Learning inquiry.** c. 2. s. 2: 127-137.

Webb, Marie, Evelyn Doman. 2016. Does The Flipped Classroom Lead to Increased Gains on Learning Outcomes in ESL/EFL Contexts? **The CATESOL Journal.** c. 28. s. 1: 39-67.

Webb, Marie, Evelyn Doman, Kerry Pusey. 2014. Flipping a Chinese University EFL Course: What Students And Teachers Think of The Model. **The Journal of Asia TEFL.** c. 11. s. 4: 53-87.

Wenden, Anita. 1991.**Learner Strategies For Learner Autonomy**. London: Prentice Hall.

W3Techs. [08.05.2016].Usage of Content Languages For WebsitesHttp://W3techs. Com/Technologies/Overview/Content_Language/All.

Wilson, Stephanie Gray. 2013. The Flipped Class: A Method to Address The Challenges of An Undergraduate Statistics Course. **Teaching of Psychology.** c. 40. s. 3: 193-199.

Yanar, Burcu Hancı, Nilay T. Bümen. 2012. İngilizce İle İlgili Özyeterlik İnancı Ölçeğinin Geliştirilmesi. **Kastamonu Eğitim Dergisi.** c. 20. s. 1: 97-110.

Yavuz, Mehmet. 2016. Ortaöğretim Düzeyinde Ters Yüz Sınıf Uygulamalarının Akademik Başarı Üzerine Etkisi ve Öğrenci Deneyimlerinin İncelenmesi. Yüksek Lisans Tezi. Atatürk Üniversitesi Eğitim Bilimleri Enstitüsü

Yıldırım, Ali, Hasan Şimşek. 2006. **Sosyal Bilimlerde Nitel Araştırma Yöntemleri.** Ankara: Seçkin Yayıncılık.

Yiğit, Raziye Demiralay. 2014. Evde Ders Okulda Ödev Modelinin Benimsenmesi Sürecinin Yeniliğin Yayılımı Kuramı Çerçevesinde İncelenmesi. Doktora Tezi. Gazi Üniversitesi Eğitim Bilimleri Enstitüsü.

Yu, Zhonggen, Qin Zhu. 2016. Comparison of Satisfaction in EFL Flipped and Traditional Classrooms.**2nd International Conference on Humanities and Social Science Research:** 326-330.

Yurtseven, Nihal, Bülent Alcı, Hakan Karataş. 2014. Factors affecting academic performance in EFL context: A modeling study. **The International Journal of Research in Teacher Education**. c. 5. s. 1: 13-23.

Yurtseven, Nihal, Sertel Altun, Hasan Aydın. 2015. An Analysis on Motivational Beliefs ofPreparatory Class About Learning English. **International Journal of Educational Researchers.** c. 6. s. 1: 13-30.

Zainuddin, Zamzami, Siti Hajar Halili. 2016. Flipped Classroom Research and Trends From Different Fields of Study. **The International Review of Research in Open and Distributed Learning.** c. 17. s. 3: 313-340.

Zappe, Sarah, Robert Leicht, John Messner, Thomas Litzinger, Hyeon Woo Lee. 2009. "Flipping" The Classroom to Explore Active Learning in A Large Undergraduate Course. **American Society for Engineering Education:** 14-35

Zeren, Merve Görkem. 2016. The Flipped Geography Lecture. **Marmara Coğrafya Dergisi**. c. 33: 25-57.

APPENDICES

Appendix 1. EFL Achievement Test

Sevgili Öğrenciler;

Bu test İngilizcenin beş farklı boyutundaki başarınızı ölçmek üzere hazırlanmıştır. İki oturumda tamamlanacak olan testte Listening (Dinleme), Grammar (Dilbilgisi), Reading (Okuma), Vocabulary (Kelime) ve Writing (Yazma) olmak üzere beş farklı bölüm bulunmaktadır. İlk oturumda bu test kitapçığındaki ilk dört bölümdeki bütün soruları cevaplamanız için 60 dakikanız bulunmaktadır. İkinci oturumdaysa son bölüm olan Writing (Yazma) bölümünü tamamlamanız için 30 dakikanız olacaktır.

Testin sonuçları sizlere daha etkili ve anlaşılır bir İngilizce dersinin geliştirilmesine katkıda bulunabileceğinden önemlidir. Bu bakımdan, lütfen tüm soruları dikkatli bir şekilde yapmaya çalışınız.

Katılımlarınız için teşekkürler

Orhan İYİTOĞLU

Yıldız Teknik Üniversitesi

Eğitim Programları ve Öğretim Doktora Öğrencisi

Adınız:

Soyadınız:

Cinsiyetiniz : (   ) Bayan   (   ) Bay

Lisans Bölümünüz:

A) LISTENING (20 PTS.)

I. Listen to the interview with Mr Smith about his lifestyle and mark the sentences True (T) or False (F). (5x2= 10 pts.)

  1. Mr Smith gets up at 6.30 a.m.   _____

  2. Mr Smith goes to work by train.   _____

  3. Mr Smith has lunch in a restaurant.   _____

  4. Mr Smith usually finishes work at 7.30 p.m. _____

  5. Mr Smith often goes out in the evening.   _____

II. Listen to five conversations and complete the following sentences (5x2= 10 pts.)

  1) Dan wants to have

  _____

2) Sue goes to bed after she

  _____

3) Karen goes to the gym

  _____

4) Jason's birthday is on

  _____

5) During the week Polly

  _____

B) GRAMMAR (20 PTS.)

I. Circle the correct answer. (8x1= 8 pts.)

1) _____ are her books.

A. This   **B.** That   **C.** Those   **D.** It

2) This isn't Tim's camera. It's _____

  A. your   **B.** my   **C.** our   **D.** mine

3) Is there _____ cheese in the bridge?

  A. any   **B.** a   **C.** an   **D.** some

4) Kim studies _____ of all the students in my class.

  A. hard   **B.** the hardest   **C.** harder   **D.** more hard

5) Bill and Mark _____ from 10 a.m. to 6 p.m. every day.

A. work   **B.** works   **C.** worked   **D.** are going to work

6) ___Sue ___ to study a foreign language?

  A. Do…have **B.** Does…has   **C.** Does…have   **D.** Did ... has

7) I didn't like the film _____ I turned off the TV.

  A. but   **B.** because   **C.** moreover   **D.** so

8) I _____ go to school because I _____ ill.

  A. can / am   **B.** couldn't / was   **C.** could / was   **D.** didn't / am

II. Put the verbs in brackets into the correct form. (8x1= 8 pts.)

Example:

A: Did *you go* (you/go) to Spain for your summer holiday?

B: No, I didn't. I *went* (go) to Portugal.

1) A: _____ (she/like) reading books?

B: Yes, she does. She _____(love) reading books.

2) A: _____ (you/play) football yesterday?

B: No, I _____ (visit) my grandparents.

3) A: _____ (they/wash) the car at the moment?

B: No, they _____ (listen) to the news.

4) A: I _____ (go) to the theatre tonight.

B: What _____ (you/wear)?

III. Rewrite the following sentences using the words below. (2x2= 4 pts.)

1) Brain is more successful in history than John.

As…….. as

_____

2) Although he studied hard, he couldn't pass the exam.

In spite of

_____

C) READING (20 PTS.)

Inverness

Inverness is a small town situated in beautiful countryside in the north-east of Scotland. It is an important centre for visits to many parts of the Highlands, which is the famous mountainous area in Scotland. Inverness is a small town but it has got an excellent transport system. There are trains to all the main cities of Scotland (Inverness - Aberdeen 1.5 hours, Inverness – Dundee 3 hours, Inverness – Glasgow 4 hours, Inverness – Edinburgh 4.5 hours). Dalcross Airport (about 8 miles east of the town) offers flights to many parts of the country. The buses leaving from Faraline Park serve the town and the surrounding area. Many of the big banks have branches in Inverness so changing money is not a problem for tourists. As well as these essential services, the town has cultural and entertainment facilities. There are theatres and cinemas, and some hotels organize Highland entertainment during the summer months.

Inverness is an old town, but there are not many buildings left from the old times. However, there are still some buildings in the town which are very old, such as the houses on Douglas Row. And Church Street has some interesting old churches.

Inverness is situated at the northern end of the Caledonian Canal. The canal was opened in 1822. Sailing boats used the canal to travel between the Irish Sea and the North Sea. The canal provided a safe and fast journey for them. When steamships became popular instead of sailing boats, people stopped using the canal for transport because these steamships were too large and could not travel down the canal. Nowadays, there are sightseeing trips for tourists down the canal.

There are many good eating places in Inverness. There are plenty of restaurants and hotels which offer meals to non-residents. They have delicious local dishes at cheap prices; try some 'haggis' and a glass of malt whisky before you leave. There are also all kinds of sea food.

I. Mark the following sentences as True (T) or False (F). (5x1= 5 pts.)

1) You can see old buildings on Douglas Row and Church Street. _____

2) People have difficulty in finding a place to change money in Inverness. _____

3) The transportation to the main cities of Scotland from Inverness is not so easy.

_____

4) Due to the inability of steamships to travel down the canal, people stopped
using the Caledonian Canal for transportation. _____

5) The writer suggests that visitors to Inverness try some of the local food. _____

II. Which of the following information can you find in the text? Put a tick (✓) in the blank of the item if you think it exists in the text or a cross (x) if you think it does not exist in the text. (5x1= 5 pts)

___ a) what the weather is like   ___ d) what to eat and drink

___ b) where Inverness is   ___ e) population of Inverness

___ c) hotel prices

III. Read the text and answer the questions. (5x2=10 pts.)

1) What is Highlands?

_____

2) What are the essential services in Inverness according to the text?

_____

3) You can travel to several Scottish cities from Inverness by train. Write the
names of these cities.

_____

4) When did people stop using the canal for transportation?

_____

5) What kind of meals do the restaurants and hotels in Inverness offer to non-
residents? (Give one example)

_____

D. VOCABULARY (20 PTS.)

I. Complete the sentences with the suitable words in the box. (8x1= 8 pts.)

> **angry   cleverest   mixture   join**
>
> **band   proud of   attic   chance**

1) Can I _____ the football club?
2) Don't miss the _____ to visit Paris.
3) John is very successful in his lessons. His parents are very _____ him.
4) There is a(n) _____ in your house. We use it as a study place.
5) The Beatles is a(n) _____ from England. I like listening to them.
6) Amy is the _____ of all my friends.
7) Julia is _____ because her children broke the window.
8) This music is a(n) _____ of rock and reggae.

II. Look at the word in brackets following each sentence. Complete each sentence by writing the correct form of the word in the gap. (6x2=12 pts.)

1) None of his _____ ever worked. (invent)
2) They waited _____ for the news. (anxious)
3) The house had large rooms and was very _____. (comfort)
4) The teacher warned him about his _____. (behave)
5) He is one of the most _____ members of the team. (create)
6) Nobody really believed his_____. (explain)

E. WRITING (20 PTS.)

I. **Write a letter introducing yourself and your family to your pen-friend using the plan below. (20 pts.)**

Plan
Introduction
(Paragraph 1) (your full name, age, what you look like, where you live)
Main Body
(Paragraph 2) (your family: names, ages, jobs, what they look like, what they like)
Conclusion
(Paragraph 3) (ask your friend to write back)

Dear ….,

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Appendix 2. Detailed Scoring Key for Achievement Test

A) LISTENING (20 PTS.)

I.   Listen to the interview with Mr Smith about his lifestyle and mark the sentences True (T) or False (F). (5x2= 10 pts.)

1. Mr Smith gets up at 6.30 a.m.   ___*F*___

2. Mr Smith goes to work by train.   ___*T*___

3. Mr Smith has lunch in a restaurant.   ___*F*___

4. Mr Smith usually finishes work at 7.30 p.m.   ___*F*___

5. Mr Smith often goes out in the evening.   ___*F*___

**II. Listen to five conversations and complete the following sentences.
   (5x2= 10 pts.)**

1)   Dan wants to have *__a cup of coffee and a toast_____*

2) Sue goes to bed after she *__watches TV / has a shower_____*

3) Karen goes to the gym *__once every two weeks / twice a month_____*

4) Jason's birthday is on *__the thirteenth of June_____*

5) During the week Polly *__gets up at about eight_____*

- In exercise II, no points will be deducted for errors of grammar and spelling as long as it is clear that the correct response was intended by the student.

B) GRAMMAR (20 PTS.)

I. **Circle the letter of correct answer. (8x1= 8 pts.)**

1. _____ are her books.
   a. This   **b.** That   **c.** Those   **d.** It

2. This isn't Tim's camera. It's _____
   a. your   **b.** my   **c.** our   **d.** mine

3. Is there _____ cheese in the bridge?
   a. any   **b.** a   **c.** an   **d.** some

4. Kim studies _____ of all the students in my class.
   a. hard   **b.** the hardest   **c.** harder   **d.** more hard

5. Bill and Mark _____ from 10 a.m. to 6 p.m. everyday.
   a. work   **b.** works   **c.** worked   **d.** are going to work

6. ___Sue ___ to study a foreign language?
   a. Do…have   **b.** Does…has   **c.** Does…have   **d.** Did ... has

7. I didn't like the film _____ I turned off the TV.
   a. but   **b.** because   **c.** moreover   **d.** so

8. I _____ go to school because I ____ ill.
   a. can / am   **b.** couldn't / was   **c.** could / was   **d.** didn't / am

II. **Put the verbs in brackets into the correct form. (8x1= 8 pts.)**

Example:

A: Did _*you go*_ (you/go) to Spain for your summer holiday?

B: No, I didn't. I _*went*_ (go) to Portugal.

1) A: _*Does she*_ (she/like) reading books?

   B: Yes, she does. She _*loves*_ (love) reading books.

293

2) A: ___*Did you*___ (you/play) football yesterday?

   B: No, I ___*visited*___ (visit) my grandparents.

3) A: ___*Are they*___ (they/wash) the car at the moment?

   B: No, they ___*are listening*___ (listen) to the news.

4) A: I ___*am going*___ (go) to the theatre tonight.

   B: What ___*are you going to wear*___ (you/wear)?

- Nothing will be deducted for non-grammatical errors or errors in grammar which is not tested by the item. As an example, "*visitd*" will be accepted for "*visited*", without penalty. But no point will be awarded for *"love"* instead of *"loves".*

III. Rewrite the following sentences using the words below. (2x2= 4 pts.)

1) Brain is more successful in history than John.

As……..as ___*John is not as successful as Brain in history* .___

2) Although he studied hard, he couldn't pass the exam.

In spite of ___*In spite of studying hard, he couldn't pass the exam.*___

- Since students are expected to write a grammatical sentence equivalent in meaning to the one that is given, points will be deducted for the errors of the grammatical structure demonstrated by the given part of the paraphrase. Again, nothing will be deducted for non-grammatical errors or for tiny errors of grammar which is not being tested by the test item. As an example, the correct answer of 1st question is "*John is not as successful as Brain in history".* Nothing will be deducted when a student answers this question as follow: *"John is not as successful as Brain on history".* But point(s) will be deducted when the elicited answer of the 2nd question is *"In spite of study hard, he couldn't pass the exam."*

294

C) READING (20 PTS.)

I. Mark the following sentences as True (T) or False (F). (5x1=5 pts.)

1)__*T*___

2)__*F*___

3)__*F*___

4)__*T*___

5)__*T*___

II. Which of the following information can you find in the text? Put a tick (√) in the blank of the item if you think it exists in the text or a cross (x) if you think it does not exist in the text. (5x1= 5 pts)

__**X**__ **a)** what the weather is like   __√__ **d)** what to eat and drink

__√__ **b)** where Inverness is   __**X**__ **e)** population of Inverness

__**X**__ **c)** hotel prices

III.   Read the text and answer the questions. (5x2=10 pts.)

1) *The Highlands is the famous mountainous area in the Scotland.*

2) a) *It has got an excellent transport system and the Dalcross Airport offers flights to many parts of the country, etc.*

   b) *There are many banks and changing money is not a problem.*

3) *Aberdeen, Dundee, Glasgow and Edinburgh.*

4) *When steamships became popular instead of sailing boats, people stopped using thecanal for transport.*

5) *They offer delicious local dishes at cheap prices. Haggis is one example of this.*

295

- Since it will make the measurement of reading ability less accurate, errors of grammar, spelling or punctuation will not be penalized provided that the student has performed the reading task which the item sets. In the 2nd and 5th questions, students can give some other examples from the text different than the ones written above.

D. VOCABULARY (20 PTS.)

I. Complete the sentences with the suitable words in the box. (8x1= 8 pts.)

> angry   cleverest   mixture   join
>
> band   proud of   attic   chance

1) Can I _join_ the football club?
2) Don't miss the _chance_ to visit Paris.
3) John is very successful in his lessons. His parents are very _proud of_ him.
4) There is a(n) _attic_ in your house. We use it as a study place.
5) The Beatles is a(n) _band_ from England. I like listening to them.
6) Amy is the _cleverest_ of all my friends.
7) Julia is _angry_ because her children broke the window.
8) This music is a(n) _mixture_ of rock and reggae.

II. Look at the word in brackets following each sentence. Complete each sentence by writing the correct form of the word in the gap. (6x2=12 pts.)

1) None of his _inventions_ ever worked. (invent)
2) They waited _anxiously_ for the news. (anxious)
3) The house had large rooms and was very _comfortable_. (comfort)
4) The teacher warned him about his _behaviours_. (behave)
5) He is one of the most _creative_ members of the team. (create)
6) Nobody really believed his _explanation_. (explain)

- In this part, nothing will be deducted for grammatical errors or for errors in anything which is not being tested by the item. As an example, in 1st and 4th

questions, students will not be penalized for a missing plural –s. But students will be penalized for some certain spelling mistakes. To make it more clear, *"anxiousl"* is not accepted instead of *"anxiously"*.

E. WRITING (20 PTS.)

Grammar– range, adequacyand accuracy

__5. Very few errors of grammar.

    Mastery of target grammatical structures.

__4. Some errors of grammar or word order which do not interfere with comprehension.

__3. Fairly frequent errors of grammar or word order which affect the meaning of what is written.

__2. More frequent errors of grammar which causes the test reader to rely on his/her own interpretation.

__1. Very frequent and severe errors of grammar which makes comprehension virtually impossible.

Vocabulary - range, adequacy and accuracy

__5. Use of wide range of vocabulary and even idioms appropriate to the task.

__4. Minor inappropriacies in use of vocabulary.

__3. Fairly frequent use of wrong and inappropriate words.

__2. Use of limited vocabulary

   Frequent errors in use of vocabulary which causes the test reader to rely on his/her own interpretation.

__1. Use of very poor range of vocabulary.

    Extreme vocabulary limitations which make it impossible to comprehend.

Content – relevance and adequacy

__5. Writing fully relevant and adequate to the task.

Uses of well developed and supported ideas.

__4. Generally relevant writing with some ideas needed to be developed a bit more.

__3. A mostly relevant letter with some repetition which are limited in details.

__2. A nearly relevant letter with more repetition.

Uses of ideas which are often difficult to follow.

__1. A letter which is mostly irrelevant with not fulfilled purpose which makes comprehension virtually impossible.

Organization - form

__5. Highly organized; clear progression of ideas into well developed, supported and linked ideas.

__4. Generally coherent with some inadequate links among the ideas

__3. Fairly frequent lack of organization which requires re-reading of the letter for the clarification of ideas.

__2. Little or no attempt at connectivity which makes it difficult for the reader to follow.

__1. Severe lack of organization as a result of which communication is seriously impaired.

SCORE: Grammar: **_5_** + Vocabulary: **_5_** + Content: **_5_** + Organization: **_5_**

Appendix 3. Stages of Test Construction

I. Statement of the problem

In order to learn what progress is being made during the English course in a prep class at Gebze Technical University, a progress test should be administered before the course comes to an end. This test is sensitive enough to gain diagnostic information about students' progress in listening, reading, writing, grammar, and vocabulary. By means of this test, we will learn how much students have learnt and whether they are good at using these skills or not; therefore, the results of the test should be as detailed and accurate as possible and by looking at these results their teachers will be able to make necessary changes on the course. Backwash is a serious consideration. The test must therefore include tasks and topics with which the students are already familiar and which will encourage the practice of listening, reading, writing skills and other skills. There is no serious constraint on the administration and the construction of the test apart from the fact that the test is long since many abilities are required to be tested. However, some parts of the test can be administered separately in order to solve this problem.

II. Specifications for the test

a) Operations:

In Part A, listening is the main skill.

Task I. This is a closed ended task since there is no production. Students will recognize the sentences and identify them as true or false. Listening for specific information to distinguish is the underlying skill and the cognitive domain is to analyze.

Task II. This task is a limited response one and students are required to find relevant information on the cassette and complete the guided short answers. This task entails students to recognize, recall, and select the relevant information; therefore, the cognitive domain is to understand.

In Part B, Grammar is the main skill.

Task I. This is a closed ended task. The sub-skill is recognizing and applying the correct grammatical structure. Students are required to choose the suitable grammatical structures for the sentences. Therefore, the cognitive domains on task are to understand and apply.

Task II. This task is a limited response one since students, based on rules, are required to change the forms of the verbs given for each sentence in order to complete them; as a result, the cognitive domain is to apply. The sub-skill is forming verbs according to the tenses of the sentences in which they are used.

Task III. This task is a limited response one. The sub-skill is to rewrite a grammatical sentence equivalent in meaning to the one that is given. Students are required to use the structures that they have already learnt in different contexts. The task requires the test takers to rewrite, manipulate and use the given structures in a different sentence; therefore, the cognitive domain is to apply.

In Part C, Reading is the main skill.

Task I. This task is a closed ended one. The sub-skill is scanning text for relevant specific information. The cognitive domain is to analyze since students are required to distinguish the sentences and identify them as true or false.

Task II. This one is a limited response task. The sub-skills are scanning the text for detailed information and making inferences from the clues given in the texts. The cognitive domain is to analyze since students are required to recognize whether some specific information exists in the text.

Task III. This is a limited response ended task. Students are required to answer some comprehension questions about the text. The cognitive domain is to understand and the sub-skill is scanning the text for specific information.

In Part D, Vocabulary is the main skill.

Task I. The task is a closed ended one. The sub-skill is choosing the right word for each sentence from the box. Students are required to understand the meaning of the sentences and words; therefore, cognitive domain is to understand.

Task II. This is a limited response task. The sub-skills are choosing the right words for each sentence and using the words in the correct form which suits the sentence

best. Students are required to modify the words given with the sentences and the cognitive domain is to apply.

In Part E, Writing is the main skill.

Task I. This is an open ended task. The sub-skills are introducing themselves and their families. The cognitive domain is to apply since it is restricted with the instructions. Students are required to use a variety of vocabulary for making description and have the necessary skills for organizing a letter for introducing themselves and their families.

b) Types of the text:

Part A, In this part, there are two parts. In the first part, there is a conversation between two people who talk about typical life in London. In the second part, there are five separate dialogues between different pairs. They are authentic-like conversations.

Part B, In this part isolated items are used. There is no specific context.

Part C, The passage in this part is about a small town called Inverness and it is taken from a coursebook.

Part D, Isolated items are used again in this part.

Part E, The text is a letter of introduction. Students are guided by a plan of the letter and this plan was extracted from a coursebook called "Click on 1".

c) Addressees:

The students are in Prep School at Gebze Technical University. They are in a mixed-ability class; therefore, their levels are mostly elementary and little to false pre-intermediate.

d) Topics:

The students' level and interests are considered important. As a result, a lot of effort has gone into making the topics as suitable as possible for the students. The listening section is composed of typical London lifestyle and real-life dialogues. The reading passage is about a small town called Inverness. The writing section is concerned with a letter about introducing yourself and your family.

e) Format and timing:

Part A Listening:

<u>Task I.</u> A true-false task with 5 items and it is closed ended. The conversation lasts 5 minutes. The students will listen to the conversation twice.

> *e.g. Listen and mark the sentences as True or False.*
>
> *American families often move house _____*
>
> Time: This task with five items will take 5 minutes.

<u>Task II.</u> A completion task with 5 items and it is a limited-response activity. Students are supposed to complete the items when listening to the conversation.

> *e.g. Listen again and complete the following sentences*
>
> *Dan wants to have _____*
>
> Time: 8 minutes will be enough for this activity.

Part B Grammar:

<u>Task I.</u> A multiple-choice task with 8 items.

> *e.g. Circle the letter of the correct answer.*
>
> *This isn't Tim's camera. It's _____*
>
> *A. your   B. my   C. our   D. mine*
>
> Time: It will take 6 minutes since students are only expected to circle the letter of correct response.

<u>Task II.</u> A formation task with 4 items.

> *e.g. Put the verbs in brackets into the correct form.*
>
> *She_____(go) to the theatre tonight.*
>
> Time: 6 minutes will be enough.

<u>Task III.</u> A paraphrase task with 2 items.

> *e.g. Rewrite the following sentences using the words below.*
>
> *Brain is more successful in history than John.*

As…..... as _____

    Time: 4 minutes will be enough again.

Part C Reading:

Task I. There is a reading passage with 4 paragraphs and approximately 400 words. First task is a true-false one with 5 items in order in which relevant information appears in the text.

    e.g. Mark the following sentences as True (T) or False (F).

    You can see old buildings on Douglas Row and Church

    Street _____

    Time: This activity will take 5 minutes.

Task II. A task with 5 items about finding the relevant information in the passage.

    e.g. Which of the following information can you find in the text?

        _____ Hotel prices

    Time: It will take 10 minutes.

Task III. A task with 5 comprehension questions.

    e.g. Read the text and answer the questions.

    When did people stop using the canal for transportation?

    Time: 6 minutes will be enough for the completion of this activity.

Part D Vocabulary:

Task I. A fill-in-the blank task with 8 items.

    e.g. Complete the sentences with the suitable words in the box.

    Can I _____ the football club?

    Time: 4 minutes

Task II. A word formation task with 6 items.

> *e.g. Complete each sentences by writing the correct form of the word in the gap.*
>
> *None of his _____ ever worked. (invent)*
>
> Time: 6 minutes

Part E Writing:

Task. Writing a letter according to the plan on the last page.

> *e.g. Write a letter introducing yourself and your family to your pen-friend using the plan below.*
>
> Time: 30 minutes

f) Criterial levels of performance:

It is possible for the students to complete all of the tasks successfully since the students are already familiar with such kinds of tasks, topics, structures and items. The total score of the test is 100. A score of 45 is sufficient for passing the exam. However, if a student can get a score of 85, he/she is considered to be successful in terms of the course's objectives.

III. Comments on sampling, item writing and moderation, writing scoring key and its moderation

In this test, a variety of items, techniques and tasks were used in order to test students' achievement in listening, reading, grammar, vocabulary and writing skills. To achieve content validity and beneficial backwash on both teacher's teaching procedures and students' learning process, we have tried to choose widely from the whole area of content. Therefore, the test we have prepared, we can claim that, includes representative samples of the skills that the students have acquired during the course. Since students are already familiar with format and testing techniques, which means that it does not put extra demands on them; therefore, we can argue that the reliability of the test will increase.

In the item writing procedure, many items were written, yet some items were rejected and others were reworked. In order to achieve a desirable quality in item writing, we consulted the English teachers at the practice school and we modified some items

and instructions according to the advice that we received. For instance, the reading passage was too difficult for the students and it was simplified according to the students' level. Moreover, due to the constraint of time, the test constructer should make sure that the test is not too long. For that reason, some items were deleted. As it is clear, considerable time was spent for moderation and rewriting of items.

The scoring key was prepared as detailed as possible in order to attain objective scoring. To evaluate students' levels of success in each skill, nothing was deducted for the errors of the skill outside the confines of what the item tested. After consulting the English teachers at the practice school, we made necessary changes in the scoring key. For incomplete answers, we tried to give partial scores. Revised version of the scoring key was presented to the teachers and they accepted it as a satisfactory one.

Appendix 4.Table of Specifications for the University Prep EFL Achievement Test

Skills	Objectives & Functions	Understand	Apply	Analyze	No of Items	Points	%
Listening	Students distinguish true and false sentences in an audio text.			X	5(A.I.)	10	20
Listening	Students complete the blanks in an audio text.	X			5(A.II.)	10	20
Grammar	Students identify the correct form of the pronouns in a sentence.	X			2(B.I. 1,2)	2	20
Grammar	Students use the correct quantifiers with countable and uncountable nouns.		X		1(B.I. 3)	1	20
Grammar	Students identify correct form of the adjectives to make comparison.	X			1(B.I. 4)	1	20
Grammar	Students identify the correct tense to complete the sentences.	X			3 (B.I. 5,6,8)	3	20
Grammar	Students combine sentences by using correct conjunctions.		X		1(B.I. 7)	1	20
Grammar	Students identify the tenses to complete the sentences with the correct form of the verbs.		X		8(B.II.)	8	20
Grammar	Students rewrite the sentences by using the given structures and words.		X		2(B.III.)	4	20
Reading	Students distinguish true and false sentences in a text.			X	5(C.I)	5	20
Reading	Students make inferences from a text by using the given clues.			X	5(C.II.)	5	20
Reading	Students answer wh- questions about a given text.	X			5(C.III)	10	20
Vocabulary	Students identify the best vocabulary item to complete a sentence.	X			8(D.I.)	8	20
Vocabulary	Students modify the words into correct form to complete the sentences.		X		6(D.II.)	12	20
Writing	Students write a descriptive letter about oneself and the family following the instructions.		X		1(E.I.)	20	20

Appendix 5. Kazanım Değerlendirme Formu

Sayın Değerlendirmeci;

İngilizce Hazırlık öğrencilerinin İngilizcenin beş farklı boyutundaki başarısını ölçmek üzere bir başarı testi geliştirilecektir. İki oturumda tamamlanması planlanan bu testte Listening (Dinleme), Grammar (Dilbilgisi), Reading (Okuma), Vocabulary (Kelime) ve Writing (Yazma) olmak üzere beş farklı bölüm bulunacaktır.

Bu testten elde edilecek veriler bilimsel bir çalışmaya ışık tutacağından sizden aşağıda dönem boyunca öğrencilerinizin kazanması gereken kazanımlardan kritik olarak belirlediklerimizi değerlendirmenizi istemekteyim. Böylece, öğrencilerin dönem kazanması gereken en temel kazanımları yoklayan kapsamlı bir test geliştirmemize katkıda bulunacaksınız. Bu doğrultuda, sizi en iyi yansıtan seçeneği '1, 2, 3, 4, 5' kutularından birine (X) koyarak belirtiniz. Bu derecelendirmede her rakam aşağıda belirtilen ifadeleri karşılamaktadır:

1: Tamamen Geçersiz

2: Geçersiz

3: Kararsız

4: Geçerli

5: Tamamen Geçerli

Katkılarınızdan dolayı teşekkür ederim.

Orhan İYİTOGLU

Yıldız Teknik Üniversitesi

Eğitim Programları ve Öğretim Doktora

Öğrencisi

orhanmka@hotmail.com

Skills	Objectives & Functions	Tamamen Geçersiz	Geçersiz	Kararsız	Geçerli	Tamamen Geçerli
		1	2	3	4	5
Listening	Students distinguish true and false sentences in an audio text.	1	2	3	4	5
	Students complete the blanks in an audio text.	1	2	3	4	5
Grammar	Students identify the correct form of the pronouns in a sentence.	1	2	3	4	5
	Students use the correct quantifiers with countable and uncountable nouns.	1	2	3	4	5
	Students identify correct form of the adjectives to make comparison.	1	2	3	4	5
	Students identify the correct tense to complete the sentences.	1	2	3	4	5
	Students combine sentences by using correct conjunctions.	1	2	3	4	5
	Students complete the sentences with the correct form of the verbs.	1	2	3	4	5
	Students rewrite the sentences by using the given structures and words.	1	2	3	4	5
Reading	Students distinguish true and false sentences in a text.	1	2	3	4	5
	Students make inferences from a text by using the given clues.	1	2	3	4	5
	Students answer wh- questions about a given text.	1	2	3	4	5
Vocabulary	Students identify the best vocabulary item to complete a sentence.	1	2	3	4	5
	Students modify the words into correct form to complete the sentences.	1	2	3	4	5
Writing	Students write a descriptive letter about oneself and the family following the instructions.	1	2	3	4	5

Appendix 6. Başarı Testi Değerlendirme Formu

Sayın Değerlendirmeci;

Bu test, İngilizce Hazırlık öğrencilerinin İngilizcenin beş farklı boyutundaki başarısını ölçmek üzere hazırlanmıştır. İki oturumda tamamlanacak olan testte Listening (Dinleme), Grammar (Dilbilgisi), Reading (Okuma), Vocabulary (Kelime) ve Writing (Yazma) olmak üzere beş farklı bölüm bulunmaktadır.

Bu testten elde edilecek veriler bilimsel bir çalışmaya ışık tutacağından sizden ekte gönderilen belirtke tablosu ve başarı testini dikkate alarak ilgili testin kapsam ve görünüş geçerliğini değerlendirmenizi istemekteyim. Bu doğrultuda, sizi en iyi yansıtan seçeneği '1, 2, 3, 4, 5' kutularından birine (X) koyarak belirtiniz. Bu derecelendirmede her rakam aşağıda belirtilen ifadeleri karşılamaktadır:

1: Tamamen Geçersiz

2: Geçersiz

3: Kararsız

4: Geçerli

5: Tamamen Geçerli

Katkılarınızdan dolayı teşekkür ederim.

<div align="center">

Orhan İYİTOGLU

Yıldız Teknik Üniversitesi

Eğitim Programları ve Öğretim Doktora Öğrencisi

orhanmka@hotmail.com

</div>

	Tamamen Geçersiz	Geçersiz	Kararsız	Geçerli	Tamamen Geçerli
KAPSAM GEÇERLİĞİ	1	2	3	4	5
GÖRÜNÜŞ GEÇERLİĞİ	1	2	3	4	5

Appendix 7. İngilizce ile İlgili Özyeterlik İnancı Ölçeği

Sayın katılımcı,

Bu ölçek iki bölümden oluşmaktadır. Birici bölümde sizinle ilgili bir kısım değişkenler, ikinci bölümde ise yabancı dil olarak İngilizce öğrenimi ile ilgili özyeterlik inancının belirlenmesi ile ilgili ifadelere yer verilmiştir. Bu ifadeler yabancı dil öğreniminin alt boyutları olan okuma, yazma, konuşma ve dinleme başlıkları altında gruplandırılmıştır.

Toplanacak veri bilimsel bir çalışmaya ışık tutacağından aşağıdaki ifadelerde sizi en iyi yansıtan seçeneği '1, 2, 3, 4, 5' kutularından birine (X) koyarak belirtiniz. Bu derecelendirmede her rakam aşağıda belirtilen ifadeleri karşılamaktadır:

1:Bana hiç uymuyor

2: Çok az uyuyor

3: Biraz uyuyor

4:Oldukça uyuyor

5: Bana tamamen uyuyor

Katkılarınızdan dolayı teşekkür ederim.

<div align="right">

Orhan İYİTOGLU

Yıldız Teknik Üniversitesi

Eğitim Programları ve Öğretim Doktora Öğrencisi

orhanmka@hotmail.com

</div>

1. BÖLÜM

Adınız:

Soyadınız:

Cinsiyetiniz : (   ) Bayan   (   ) Bay

Lisans Bölümünüz:

Maddeler	Bana hiç uymuyor	Çok az uyuyor	Biraz uyuyor	Oldukça uyuyor	Bana tamemen uyuyor
OKUMA	1	2	3	4	5
1. İngilizce bir metin okuduğumda anlayabilirim.	1	2	3	4	5
2. İngilizce akademik metinler okuduğumda önemli noktaları anlayabilirim.	1	2	3	4	5
3. Okuduklarımı zihnimde canlandırabilirim.	1	2	3	4	5
4. Okuduğum İngilizce metnin temasını ya da ana fikrini bulabilirim.	1	2	3	4	5
5. İngilizce bir metinle ilgili soruları cevaplayabilirim.	1	2	3	4	5
6. Okuduğum İngilizce bir metinde anlamını bilmediğim sözcükleri tahmin edebilirim.	1	2	3	4	5
7. İngilizce bir metinde aradığım bilgiyi kolaylıkla bulabilirim	1	2	3	4	5
8. İngilizce sınavlarının okuma bölümlerinde başarılı olacağıma inanıyorum	1	2	3	4	5
YAZMA	1	2	3	4	5
1. İyi bir paragraf ya da kompozisyon yazabilirim.	1	2	3	4	5
2. İngilizce bir paragraf ya da kompozisyon yazarken dilbilgisi kurallarını doğru kullanabilirim.	1	2	3	4	5
3. İngilizce bir metin yazarken noktalama işaretlerini doğru kullanabilirim.	1	2	3	4	5
4. İngilizce bir metin yazarken düşüncelerimi tam ve açık olarak ifade edebilirim.	1	2	3	4	5
5. Bir şeyi İngilizce yazamadığımda, pes etmek yerine sorunu çözmek için çaba sarf ederim.	1	2	3	4	5
6. İngilizce yazarken önemli noktaları vurgulayabilirim.	1	2	3	4	5
7. İngilizce bir metni kendi cümlelerimle yeniden yazabilirim.	1	2	3	4	5
8. Günlük yaşamda kendimi İngilizce yazılı olarak ifade edebilirim (özgeçmiş, başvuru formu, şikâyet mektubu vb.)	1	2	3	4	5
9. İngilizce herhangi bir şey yazdıktan sonra hatalarımın farkına varabilirim.	1	2	3	4	5
10. İngilizce yazma ile ilgili verilen etkinlikleri yaparken yardıma ihtiyaç duyarım.	1	2	3	4	5

DİNLEME	1	2	3	4	5
1. İngilizce konuşulanları anlayabilirim.	1	2	3	4	5
2. Dinlediğim İngilizce konuşmanın ana fikrini çıkarabilirim.	1	2	3	4	5
3. Dinlediğim bir cümledeki duygusal vurguları anlayabilirim.	1	2	3	4	5
4. İngilizce bir konuşma dinlediğimde bilmediğim sözcüklerin anlamını tahmin edebilirim.	1	2	3	4	5
5. İngilizce bir konuşma duyduktan sonra duyduklarımla ilgili soruları cevaplayabilirim.	1	2	3	4	5
6. İngilizce televizyon kanallarını/ filmleri izlediğimde dinlediklerimi anlayabilirim.	1	2	3	4	5
7. Bir konuşma dinlediğimde resmi dil ile günlük konuşma dilini ayırt edebilirim.	1	2	3	4	5
8. İngilizce bir okuma parçasını dinlerken duyduklarımı doğru olarak yazabilirim.	1	2	3	4	5
9. İki kişi arasında geçen kısa bir İngilizce konuşmayı anlayabilirim.	1	2	3	4	5
10. İngilizce sınavlarının dinleme bölümlerinde başarılı olacağıma inanıyorum.	1	2	3	4	5
KONUŞMA	1	2	3	4	5
1. Günlük yaşamda gerekli ihtiyaçlarımı İngilizce'yi kullanarak karşılayabilirim. (Yurt dışında olduğunuzu düşünün, yer-yön bulma, alış-veriş vb.)	1	2	3	4	5
2. Bir mülakatta kendimi İngilizce olarak ifade edebilirim. (Üniversiteye giriş, iş başvurusu vb.)	1	2	3	4	5
3. Amaca ve duruma göre resmi ya da resmi olmayan bir şekilde İngilizce konuşabilirim.	1	2	3	4	5
4. İngilizce sorulan sorulara cevap verebilirim.	1	2	3	4	5
5. Karşımdaki beni anlamadığında düşüncelerimi başka şekilde ifade edebilirim.	1	2	3	4	5
6. Anadili İngilizce olan bir kişinin anlayabileceği şekilde İngilizce konuşabilirim.	1	2	3	4	5

Appendix 8. Permission to Use EFL Self-Efficacy Belief Scale
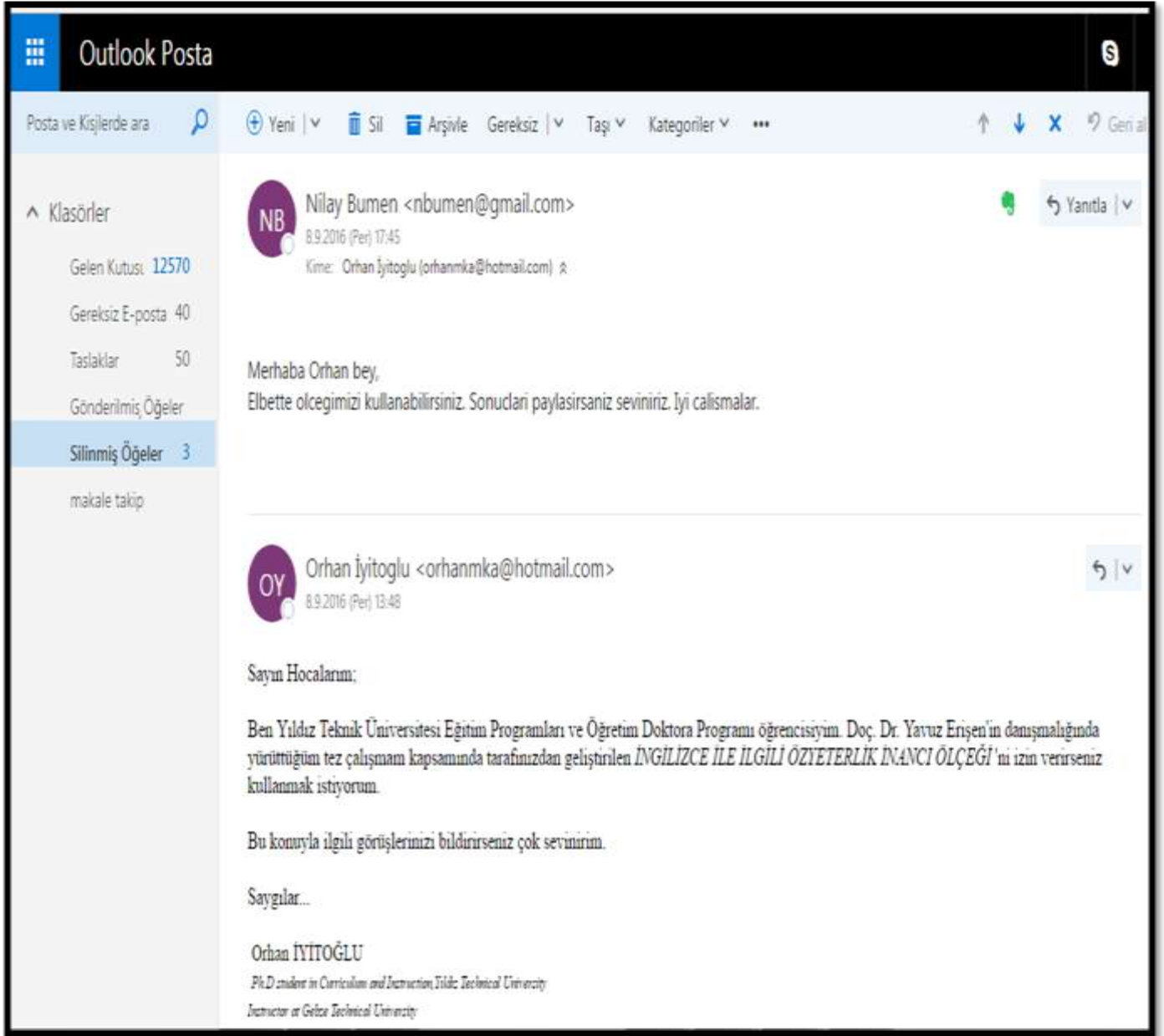
Appendix 9. İngilizce Dersine Yönelik Tutum Ölçeği

Sayın katılımcı,

Bu ölçek iki bölümden oluşmaktadır. Birici bölümde sizinle ilgili bir kısım değişkenler, ikinci bölümde ise yabancı dil olarak İngilizce öğrenimine ve dersine yönelik tutumunuzu belirlemek üzere ifadelere yer verilmiştir.

Toplanacak veri bilimsel bir çalışmaya ışık tutacağından aşağıdaki ifadelerde sizi en iyi yansıtan seçeneği '1, 2, 3, 4, 5' kutularından birine (X) koyarak belirtiniz. Bu derecelendirmede her rakam aşağıda belirtilen ifadeleri karşılamaktadır:

1: Kesinlikle katılmıyorum

2: Katılmıyorum

3: Kararsızım

4: Katılıyorum

5: Kesinlikle katılıyorum

Katkılarınızdan dolayı teşekkür ederim.

<div align="right">

Orhan İYİTOGLU

Yıldız Teknik Üniversitesi
Eğitim Programları ve Öğretim Doktora
Öğrencisi

orhanmka@hotmail.com

</div>

1. BÖLÜM

Adınız:

Soyadınız:

Cinsiyetiniz : (   ) Bayan   (   ) Bay

Lisans Bölümünüz:

Maddeler	Kesinlikle katılmıyorum	Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle katılıyorum
1. Daha fazla İngilizce bilgisi öğrenmeye meraklıyım.	1	2	3	4	5
2. İngilizce beynimizi geliştirir ve düşünmeyi öğretir.	1	2	3	4	5
3. Almam gerekenden daha fazla İngilizce dersi almak isterim.	1	2	3	4	5
4. İngilizce benim için zevkli ve güdüleyicidir.	1	2	3	4	5
5. Genellikle okulda İngilizce öğrenmekten zevk aldım.	1	2	3	4	5
6. İngilizce alıştırmalar yapmayı severim.	1	2	3	4	5
7. Eğitimim boyunca alabildiğim kadar İngilizce almaya çalışırım	1	2	3	4	5
8. İngilizce çalışırken son derece sakinim	1	2	3	4	5
9. İngilizceyi anlamaya çalışmak beni endişelendirmez.	1	2	3	4	5
10. İnsanlar için diğer dersler İngilizceden daha önemlidir.	1	2	3	4	5
11. İngilizce benim en çok korktuğum derslerden biridir.	1	2	3	4	5
12. İngilizce kendimi gergin ve rahatsız hissetmeme sebep oluyor.	1	2	3	4	5
13. İngilizceyi çok çalışmak için çok istekli değilim.	1	2	3	4	5
14. İngilizce çalışmayı nadiren severim.	1	2	3	4	5
15. İngilizce anlamsız ve sıkıcıdır.	1	2	3	4	5
16. İngilizce çok ilgi çekici bir ders değildir.	1	2	3	4	5
17. Almam gerekenden daha fazla İngilizce dersi almak istemem.	1	2	3	4	5
18. İngilizce çok değerli ve gerekli bir derstir.	1	2	3	4	5
19. İngilizce becerilerimi geliştirmek ve bu dili daha fazla öğrenmek isterim.	1	2	3	4	5

Appendix 10. Permission to Use Attitudes toward English Scale
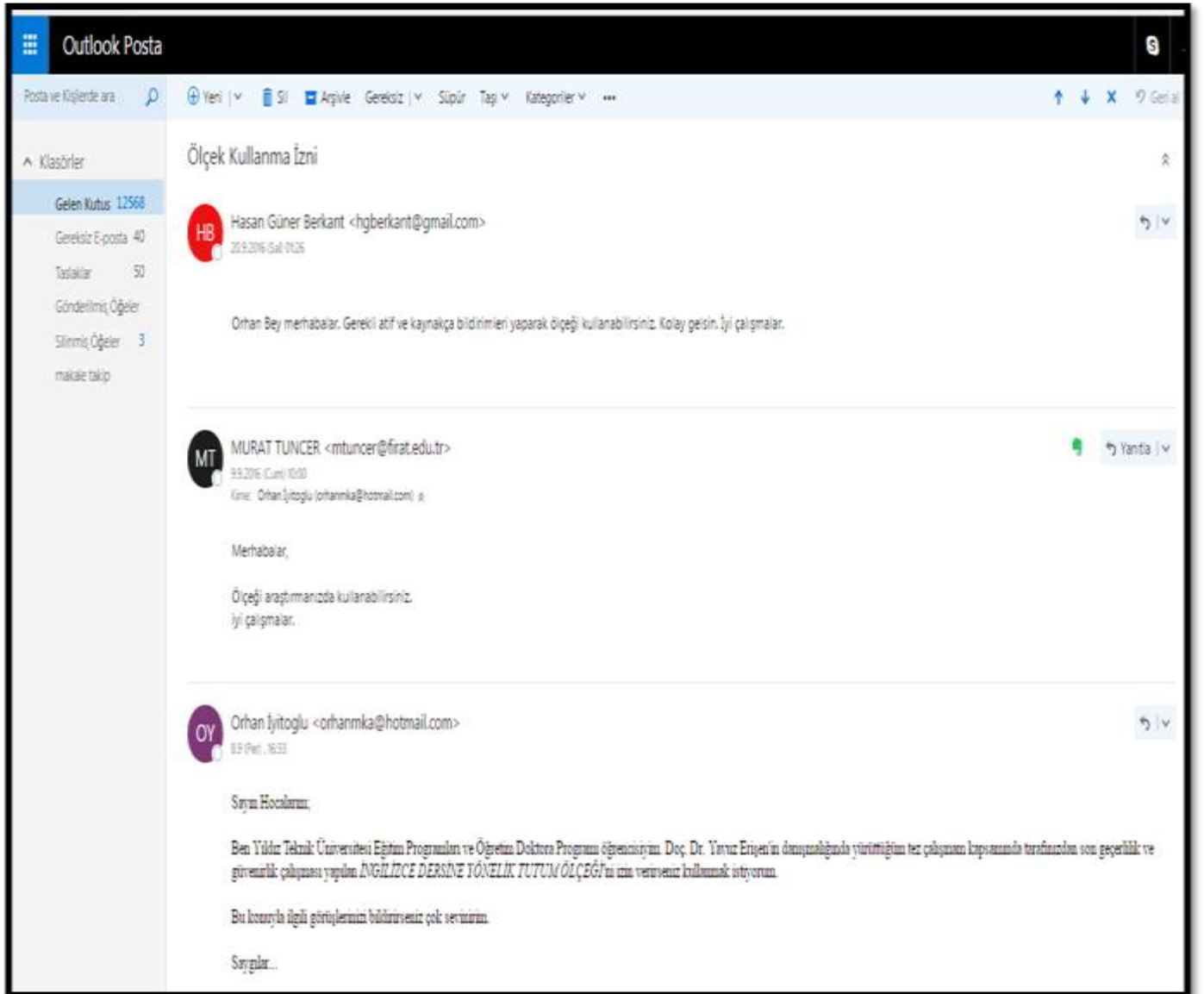
Appendix 11. Yarı Yapılandırılmış Görüşme Formu

1. Ters-yüz Sınıf Modelinin (Flipped Classroom) size sağladığı yararlar nelerdir?

2. Ters-yüz Sınıf Modeli (Flipped Classroom) uygulanırken karşılaştığınız sorunlar var mıdır? Varsa nelerdir?

3. Ters-yüz Sınıf Modeli (Flipped Classroom) ders çalışma alışkanlığınızı ve öğrenme deneyiminizi nasıl etkilemiştir?

4. Ters-yüz Sınıf Modeli (Flipped Classroom) okul dışında derse genelde harcadığınız zamanı nasıl etkilemiştir?

5. İngilizce derslerinizde bu modelin kullanılmasına devam edilmesini ister misiniz? Neden?

6. Ters-yüz Sınıf Modelinin (Flipped Classroom) başka derslerinizde de uygulanmasını ister misiniz? Neden?

7. Ters-yüz Sınıf Modeli (Flipped Classroom) İngilizce dersine karşı olan tutumunuzu nasıl etkilemiştir?

8. Ters-yüz Sınıf Modelinin (Flipped Classroom) ders dışı online ödev yapısı ve ders içi etkinlikleri İngilizce dersini daha eğlenceli mi yoksa sıkıcı mı kılmıştır? Neden?

9. Ters-yüz Sınıf Modeli (Flipped Classroom) İngilizce öğrenebileceğinize olan güven ve inancınızı nasıl etkilemiştir?

10. Son olarak eklemek istediğiniz düşünceleriniz var mı?

Appendix 12. Yarı Yapılandırılmış Görüşme Soruları Değerlendirme Formu

Sayın katılımcı,

Bu görüşme formu ters-yüz sınıf modeline (flipped classroom) göre ingilizce derslerini işleyen öğrencilerin ilgili model hakkındaki deneyim ve düşüncelerini detaylandırmak üzere hazırlanmıştır. Bu bağlamda, ilk 6 soru genel olarak ters-yüz modelin öğrencilerin ingilizce öğrenme deneyimi ve alışkanlıkları üzerindeki etkisini, bu modelin diğer derslerde de kullanılıp kullanılmamasıyla ilgili düşüncelerini sorgulamaktadır. 7 ve 8. sorular ters-yüz sınıf modelinin öğrencilerin ingilizce dersine karşı tutumunu nasıl etkilediği üzerinde dururken 9. soru modelin öğrencilerin İngilizce öğrenme ile ilgili özyeterlik inançlarına etkisini incelemektedir.

Toplanacak veri bilimsel bir çalışmaya ışık tutacağından aşağıdaki soruların ilgili yukarıda belirtilen noktaları kapsayıp kapsamadığını değerlendirmenizi istemekteyim. Bu doğrultuda her soru için sizi en iyi yansıtan seçeneği '1, 2, 3, 4, 5' kutularından birine (X) koyarak belirtiniz. Bu derecelendirmede her rakam aşağıda belirtilen ifadeleri karşılamaktadır:

1:Tamamen Geçersiz

2:Geçersiz

3:Kararsız

4:Geçerli

5:Tamamen Geçerli

Katkılarınızdan dolayı teşekkür ederim.

<div align="right">

Orhan İYİTOGLU

Yıldız Teknik Üniversitesi

Eğitim Programları ve Öğretim Doktora

Öğrencisi

orhanmka@hotmail.com

</div>

Görüşme Soruları	Tamamen Geçersiz	Geçersiz	Kararsız	Geçerli	Tamamen Geçerli
Ters-yüz Sınıf Modelinin (Flipped Classroom) size sağladığı yararlar nelerdir?	1	2	3	4	5
Ters-yüz Sınıf Modeli (Flipped Classroom) uygulanırken karşılaştığınız sorunlar var mıdır? Varsa nelerdir?	1	2	3	4	5
Ters-yüz Sınıf Modeli (Flipped Classroom) ders çalışma alışkanlığınızı ve öğrenme deneyiminizi nasıl etkilemiştir?	1	2	3	4	5
Ters-yüz Sınıf Modeli (Flipped Classroom) okul dışında derse genelde harcadığınız zamanı nasıl etkilemiştir?	1	2	3	4	5
İngilizce derslerinizde bu modelin kullanılmasına devam edilmesini ister misiniz? Neden?	1	2	3	4	5
Ters-yüz Sınıf Modelinin (Flipped Classroom) başka derslerinizde de uygulanmasını ister misiniz? Neden?	1	2	3	4	5
Ters-yüz Sınıf Modeli (Flipped Classroom) İngilizce dersine karşı olan tutumunuzu nasıl etkilemiştir?	1	2	3	4	5
Ters-yüz Sınıf Modelinin (Flipped Classroom) ders dışı online ödev yapısı ve ders içi etkinlikleri İngilizce dersini daha eğlenceli mi yoksa sıkıcı mı kılmıştır? Neden?	1	2	3	4	5
Ters-yüz Sınıf Modeli (Flipped Classroom) İngilizce öğrenebileceğinize olan güven ve inancınızı nasıl etkilemiştir?	1	2	3	4	5
Son olarak eklemek istediğiniz düşünceleriniz var mı?	1	2	3	4	5

Appendix 13. Consent Form

Başlık: "Exploring the Impact Of Flipped Classroom Model On EFL Learners' Academic Achievement, Attitudes And Self-Efficacy Beliefs: A Mixed Method Study"

Aşağıda sunulan araştırmanın çerçevesi ilgili çalışmaya katılıp katılmama konusunda vereceğiniz karara yardımcı olması hedeflenmiştir. Bu bağlamda, araştırmaya katılıp katılmamakta özgür olmakla birlikte çalışmadan herhangi bir aşamada ayrılma hakkına da sahip olduğunuzu belirtmek isterim.

Bu çalışma "Evde Ders Okulda Ödev" ya da "Ters-Yüz Sınıf" şeklinde literatürde son yıllarda yer alan bir yaklaşımın Yabancı Dil olarak İngilizce Öğrenimi'nin üzerindeki etkisini yarı deneysel bir desen ile ortaya koymayı amaçlamaktadır. Bu etkiyi ortaya koyarken, adı geçen yeni yaklaşımın Türkiye'de İngilizce başarısının yordayıcısı oldukları sıklıkla vurgulanan tutum ve özyeterlik inancının üzerideki etkisini incelemek de ayrıca amaçlanmıştır.

Bu kapsamda, ilgili çalışmada ters-yüz sınıf anlayışına yönelik olarak bir dönem boyunca yapılacak öğretimin öncesinde ve sonrasında İngilizce başarı testi, İngilizce dersine yönelik tutum ölçeği ve İngilizce ile ilgili özyeterlik inancı ölçeği uygulanacaktır. Çalışmanın sonunda ön ve son test puanları olumsuz, olumlu ve orta derecede değişen olarak aranızdan seçilen toplam 9 kişi ile görüşmeler gerçekleştirilecek ve sizin onayınız ile görüşmeler çalışma kapsamındaki yapılacak analiz çalışmaları için kayıt edilecektir.

Bu çalışmaya katılmadan önce yada ölçekleri doldururken merak ettiğiniz herşeyi sorabilirsiniz. Bu çalışma süresince vereceğiniz kişisel bilgiler ikinci yada üçüncü kişilerle paylaşılmayacaktır. Ancak doldurmanız gereken ölçeklerin ilk kısmında adınız ve soyadınız istenmektedir. Çalışmanın amacına ulaşabilmesi için bu bilgilerin birbirleriyle eşleştirilmesi gerekmektedir. Bu bakımdan bu bölümlerde gerçek adınızı ve soyadınızı paylaşabileceğiniz gibi her ölçeğe aynısını yazmak koşuluyla takma isimler de kullanabilirsiniz.

Değerli katılımlarınızla gerçekleştirilebilecek bu çalışma ile, Türkiye' de yabancı dil olarak İngilizce öğrenme ile ilgili yaygın olarak görülen başarısızlık problemi konusunda farklı kesimlerden tüm ilgililere etkili bir çözüm yolu sunulacaktır.

Lütfen bu onay formunu imzalayınız. bu imza ile çalışmanın amacı ve süreci konusunda bilgilendirildiğinizi ve çalışmaya katılmayı kabul ettiğinizi beyan ediyorsunuz. Bu formun bir kopyası saklamanız için size verilecektir.

_____   _____

İmza   Tarih

Orhan IYITOGLU

Eğitim Programları ve Öğretim Doktora Programı,

Yıldız Teknik Üniversitesi, İstanbul.

Appendix 14. Application for IRB Approval to Conduct Study

Gebze Teknik Üniversitesi

Yabancı Diller Bölümü Başkanlığı'na

07.09.2016

Kayıt no - E-11233

Yıldız Teknik Üniversitesi Eğitim Bilimleri Anabilim Dalı Eğitim Programları ve Öğretim Doktora Programı kapsamında "Exploring The Impact Of Flipped Classroom Model On EFL Learners' Academic Achievement, Attitudes And Self-Efficacy Beliefs: A Mixed Method Study" başlıklı çalışmamı 2016-2017 eğitim-öğretim yılında Gebze Teknik Üniversitesi Yabancı Diller Bölümü'nde yürütebilmek adına izninizi istemekteyim.

Gereğini saygılarımla arz ederim.

Orhan İyitoğlu

Okutman, Gebze Teknik Üniversitesi

Tel: (0) 505 861 32 61

Adres: Sultançiftliği Mah. Turhat Özal Bul. Megakent Sitesi.

B7 Blok Daire: 29 Taşdelen / İstanbul

Appendix 15. IRB Permission to Conduct Study

GEBZE
TEKNİK ÜNİVERSİTESİ

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ REKTÖRLÜĞÜ
Yabancı Diller Bölümü Başkanlığı

Sayı : 46255391-199/
Konu : Doktora çalışmanız

Sayın Orhan İyitoğlu

İlgi : 07.09.2016 tarihli dilekçeniz.

İlgi dilekçenizde belirmiş olduğunuz bilimsel çalışmanızı Bölümümüzde yürütmeniz uygun bulunmuştur.

Bilgilerinize rica ederiz.

e-imzalıdır
Prof.Dr. Erhan DEMİRBAŞ
Bölüm Başkanı V.

Adres:GEBZE TEKNİK ÜNİVERSİTESİ Cumhuriyet Mahallesi, 2254 Sokak, No: 2
41400 Gebze/KOCAELİ
Telefon:+90 262 605 (18 75) Faks+90 262 654 27 31
Elektronik Ağ:www.gtu.edu.tr

Bilgi için: Ayhan KARATAŞ
Unvanı: Şef
Tel No: +90 262 6051875

Appendix 16. Syllabus for the Experimental and Control Groups

WEEK	SUBJECT	NOTES
1 **19TH - 23RD SEPTEMBER**	Total hours: 20 ENGLISH FILE ELEMENTARY UNITS: 1ABC & 2ABC Practical English is optional	Workbook and the pack can be assigned as homework.
2 **26TH - 30TH SEPTEMBER**	Total hours: 20 ENGLISH FILE ELEMENTARY UNITS: 3ABC & 4ABC Practical English is optional	Workbook and the pack can be assigned as homework.
3 **3RD - 7TH OCTOBER**	Total hours: 20 ENGLISH FILE ELEMENTARY UNITS: 5ABC & 6ABC Practical English is optional	Workbook and the pack can be assigned as homework.
4 **10TH - 14TH OCTOBER**	Total hours: 20 ENGLISH FILE ELEMENTARY UNITS: 7ABC & 8ABC & 9A Practical English is optional	Workbook and the pack can be assigned as homework.
5 **17TH - 21ST OCTOBER**	Total hours: 20 ENGLISH FILE ELEMENTARY UNITS: 9BC & 10ABC & 11AB Practical English is optional	Workbook and the pack can be assigned as homework.
6 **24TH - 28TH OCTOBER**	Total hours: 20 ENGLISH FILE ELEMENTARY UNITS: 11C & 12ABC Practical English is optional	ELEMENTARY EXAM
7 **31ST OCTOBER - 4TH NOVEMBER**	Total hours: 20 ENGLISH FILE PRE-INTERMEDIATE UNITS: 1ABC & 2ABC Practical English is optional	Workbook and the pack can be assigned as homework.
8 **7TH - 11TH NOVEMBER**	Total hours: 20 ENGLISH FILE PRE-INTERMEDIATE UNITS: 3ABC & 4ABC Practical English is optional	Workbook and the pack can be assigned as homework.

WEEK	SUBJECT	NOTES
9 **14TH - 18TH NOVEMBER**	Total hours: 20 ENGLISH FILE PRE-INTERMEDIATE UNITS: 5ABC & 6ABC Practical English is optional	Workbook and the pack can be assigned as homework.
10 **21ST - 25TH NOVEMBER**	Total hours: 20 ENGLISH FILE PRE-INTERMEDIATE UNITS: 7ABC & 8ABC & 9A Practical English is optional	Workbook and the pack can be assigned as homework.
11 **28TH NOVEMBER - 2ND DECEMBER**	Total hours: 20 ENGLISH FILE PRE-INTERMEDIATE UNITS: 9BC & 10ABC Practical English is optional	**MIDTERM** Workbook and the pack can be assigned as homework.
12 **5TH - 9TH DECEMBER**	FOCUS 2: UNITS 1 (5 HOURS) LISTENING Q SKILLS 2: UNIT 1 (5 HOURS) GREAT WRITING 2: UNIT 1 (5 HOURS)	
13 **12TH - 16TH DECEMBER**	FOCUS 2: UNITS 2 (5 HOURS) LISTENING Q SKILLS 2: UNIT 2 (5 HOURS) GREAT WRITING 2: UNIT 2 (5 HOURS)	
14 **19TH - 23ND DECEMBER**	FOCUS 2: UNITS 3 (5 HOURS) LISTENING Q SKILLS 2: UNIT 3 (5 HOURS) GREAT WRITING 2: UNIT 3 (5 HOURS)	
15 **26TH - 30TH DECEMBER**	FOCUS 2: UNITS 4 (5 HOURS) LISTENING Q SKILLS 2: UNIT 4 (5 HOURS) GREAT WRITING 2: UNIT 4 (5 HOURS)	**LEVEL EXAM**

CURRICULUM VITAE

orhanmka@hotmail.com

PERSONAL INFORMATION

Name & Surname: Orhan İYİTOĞLU

Date of Birth: 09. 03. 1984

Place of Birth: İstanbul

EDUCATION

2013-   :   Ph.D.in Curriculum and Instruction, Yıldız Technical University, Faculty of Education, Department of Education

2008-2011:   Master in Applied Linguistics/ELT, Marmara University, Faculty of Education, Department of Applied Linguistics/ELT.

2002-2006: BA in Applied Linguistics/ELT, Marmara University, Faculty of Education, Department of Applied Linguistics/ELT.

1998-2002: Umraniye High School

PROFESSIONAL EXPERIENCE

2017-   : Instructor at Gebze Technical University

2016-2017: Executive vice Assistance at Gebze Technical University

2010-2016: Teaching in Tunç Çapa Anatolian High School

2010-2011: Term I, Deputy Pricipal in Çekmeköy Toki High School

2009-2010: Deputy Principal in Rahmi Mihriban Bedestenci Primary School

2006-2010: Teaching in Rahmi Mihriban Bedestenci Primary School