

**An Investigation into the Cross-cultural
Measurement Equivalency of the CCTDI to Assess
the Critical Thinking Dispositions of Pre-service
Teachers across Turkish and American Higher
Education Institutions**

Gökhan İskifođlu

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the Degree of

Doctor of Philosophy
in
Educational Sciences

Eastern Mediterranean University
February 2013
Gazimađusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Elvan Yılmaz
Director

I certify that this thesis satisfies the requirements as a thesis for the degree of Doctor of Philosophy in Educational Sciences.

Assist. Prof. Dr. Hüseyin Yaratın
Chair, Department of Educational Sciences

I certify that I have read this thesis and that in my opinion it is fully adequate in scope and quality as a thesis for the degree of Doctor of Philosophy in Educational Sciences.

Assist. Prof. Dr. Ali Sıdkı Ağazade
Supervisor

Examining Committee

1. Prof. Dr. A. Ata Tezbaşaran _____
2. Prof. Dr. Bekir Özer _____
3. Prof. Dr. Meral Aksu _____
4. Assist Prof. Dr. Ali Sıdkı Ağazade _____
5. Assist Prof. Dr. Hatice Nilay Hasipođlu _____

ABSTRACT

The author describes the translation, validation, and Cross-cultural applicability of a multidimensional inventory of students' evaluation of critical thinking dispositions (California Critical Thinking Disposition Inventory). The goals were to translate the CCTDI, assess its psychometric properties and examine its Cross-cultural equivalency through confirmatory factor analysis and testing its measurement invariance across American and Turkish samples. Based on the data from 583 Turkish students and 448 American students from different teacher education programs, the translated Turkish version and the original English version of CCTDI displayed positive psychometric properties, thus supported the applicability of the CCTDI in Turkish educational context. However, Cross-cultural comparison of factorial structure produced poor fit of the hypothesized multidimensional model of CCTDI to the combined sample. Further analysis, based on the modification indices, supported the use of four-factor model with reduced items for Cross-cultural comparison. In addition, possible reasons for poor model fit and noninvariance across cultural groups were outlined and discussed.

Keywords: CCTDI, critical thinking, adaptation, validation, Cross-cultural equivalency.

ÖZ

Bu çalışmanın amacı Kaliforniya Eleştirel Düşünme Eğilimi Envanterini İngilizce kaynak dilinden Türkçe hedef diline çevirmek, envanterin psikometrik özelliklerini test etmek, doğrulayıcı faktör analizi ve ölçüm değişmezliği testleriyle envanterin kültürlerarası karşılaştırmalı çalışmalar için denkliğini Amerikan ve Türk örnekleriyle değerlendirmektir. Farklı öğretmen eğitimi bölümlerinden 583 Türk öğrenci ve 448 Amerikan öğrenciden elde edilen verilerin analiz sonuçları envanterin hem çevrilen Türkçe versiyonunun hem de orijinal İngilizce versiyonunun psikometrik özelliklerinin belli kriterlerin üzerinde olduğunu ve Türkçe versiyonunun Türk kültüründe uygulanabilir olduğunu gösterdi. Ancak, kültürlerarası karşılaştırmalı çalışmalar açısından envanterin faktör yapısı incelendiğinde öne sürülen yedi-faktörlü yapının iki kültürden gelen veriyle uyummadığı görülmüştür. Ölçüm değişmezliği testi sonucunda elde edilen ileri istatistiksel sonuçlar envanterin dört-boyutlu ölçme modelinin kültürlerarası çalışmalarda kullanılabileceğini göstermiştir.

Anahtar kelimeler: CCTDI, eleştirel düşünme, uyarlama, geçerleme, kültürler arası denklik.

To my parents, Cemile and Semsettin İskifođlu

I must be blessed to have you.

ACKNOWLEDGMENTS

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First and foremost, my utmost gratitude to Assist. Prof. Dr. Ali Sıdkı Ağazade, my supervisor whose sincerity and encouragement I will never forget. Without him, this dissertation would not have been possible. Prof. Dr. Necdet Osam and Assoc. Prof. Dr. Hamit Caner have been my inspiration as I hurdle all the obstacles in my personal life and the completion of this piece of work;

I would also like to thank the dissertation examining committee members for their kind efforts in undertaking the review process and for their invaluable responses for strengthening the current dissertation study;

Assist. Prof. Dr. Hatice Nilay Hasipoğlu and Assoc. Prof. Dr. Mesude Atay, for their persistent support as my thesis monitoring committee members;

I would like to thank to my aunt, Mine Iskifoğlu for her beautiful hart and mind, wisdom, insightful attitudes, and personality. You are a great person.

I would like to provide my heartiest thanks to my sisters Ruhsan Iskifoğlu and Esra Aksoy for their invaluable academic as well as personal supports and guidance in my life. They were always ready to help me whenever I needed. You were very sensitive to my calls and became the ones who understood me well;

Berrin and Ömer Faruk Çelik, my mother and father in law, for becoming my real mother and father in life, for opening a special place in their hearts for me, for

loving me unconditionally and providing me persistent help in every segment of my life;

My grandmother, Muzaffer Meryem Erdem, for her endless prays, beliefs in my success, continuous moral supports, and positive thoughts;

My special thanks to Selim Aksoy and his family for their kind personality and sensitivity;

Bayram Güzer and İlkin Çelikli, my friends, for their real friendship and moral supports. Guys, I fell so happy and lucky to be a friend of you;

Assist. Prof. Dr. Ali H. Civelek, who was never tired in helping, guiding and suggesting me during my tough times. You are a great person;

Assist. Prof. Dr. James William Kusch, who had kind concern and consideration regarding my academic requirements;

Prof. Dr. Bekir Özer for his encouragements, positive attitudes and professional supports;

Assist. Prof. Dr. Hasan Ozder, for his beliefs in my academic achievements and moral supports throughout the process of this dissertation study;

Nilufer Engin and Atilla Engin, for their steadfast encouragement to complete this study;

Ahmet Sertcanlı, for his kind efforts in assisting me through my library research and issues associated with borrowing books.

Dr. Paula FitzGibbon, for her invaluable contributions regarding qualitative research paradigm;

Dr. Peter Facione, author of the CCTDI, for the insights and feedbacks he has shared;

Participants of this research whose names are not here in this list but did a lot to achieve this dissertation;

Especially, to Tubanur Çelik Iskifođlu, my love, my woman, my wife, my friend, my child, my dream, and my everything. I cannot express how much you have contributed to the process of this piece. However, in a special part of my life, you have raised as my sun shining through the shadows of my time. You have always been there, behind me and near me. I have always found you in my heart and found myself in yours. Thank you for resisting me, tolerating me, taking me, and thank you for loving me;

I thank each of the persons above for their invaluable contributions.

TABLE OF CONTENTS

ABSTRACT.....	iii
ÖZ.....	iv
DEDICATION.....	v
ACKNOWLEDGEMENTS.....	vi
LIST OF TABLES	xi
LIST OF FIGURES	xii
1. INTRODUCTION.....	1
1.1 Background of the Study	1
1.2 Problem to be Investigated.....	13
1.3 Purpose of the Study	15
1.4 Research Questions.....	15
1.5 Significance of the Study.....	16
1.6 Summary	18
2. LITERATURE REVIEW	20
2.1 Introduction.....	20
2.2 Conceptualization of Critical Thinking and Its Components.....	21
2.3 Nurturing Critical Thinking of Pre-service Teachers.....	25
2.4 Assessing Critical Thinking Skills and Dispositions	29
2.5 Cross-cultural Research	34
2.5.1 Approaches to Studying Culture	35
2.5.2 Equivalence.....	37
2.5.3 Bias	41
3. METHODOLOGY	43

3.1 Introduction.....	43
3.2 Research Design.....	44
3.3 Theoretical and Operational Construction of the CCTDI.....	45
3.4 Translation and Back-Translation Process.....	52
3.5 Sampling and Data Collection Procedures.....	57
3.6 Procedures for Analyzing the Psychometric Properties.....	61
3.7 Procedures for Analyzing the Cross-cultural Factorial Validity.....	65
3.8 Procedures for Analyzing the Cross-cultural Measurement Invariance.....	67
3.9 Limitations of the Study.....	69
3.10 Summary.....	70
4. RESULTS.....	72
4.1 Introduction.....	72
4.2 Linguistic Equivalency across Language Versions.....	73
4.3 Psychometric Properties of the CCTDI across Language Versions.....	78
4.4 Cross-cultural Factorial Validity of the CCTDI.....	85
4.5 Cross-cultural Measurement Invariance of the CCTDI.....	89
4.6 Critical Thinking Dispositions of Pre-service Teachers.....	93
5. DISCUSSION.....	95
6. CONCLUSION.....	101
REFERENCES.....	103
APPENDICES.....	125
Appendix A: CCTDI Authorization Translator Permission Letter.....	126
Appendix B: Translated Turkish Version of the CCTDI.....	127
Appendix C: Means, Standard Deviations, and Pearson Moment Product Correlations across Turkish and US Samples.....	133

LIST OF TABLES

Table 3.5.1. Demographic Characteristics of Participants.....	60
Table 4.2.1. Distribution of Problematic Items across the Types of Linguistic Equivalency.....	74
Table 4.3.1. Means and Standard Deviations for Turkish and American Samples.....	79
Table 4.3.2. Content Validity Indices and Alpha Coefficients for Sub-scales across Turkish and English Versions of the CCTDI.....	81
Table 4.3.3. Correlations, Paired t-tests, and Variances for Sub-scale Scores of the Turkish CCTDI across Time.....	82
Table 4.3.4. Correlations, Paired t-tests, and Variances for Sub-scale Scores of the English CCTDI across Time.....	83
Table 4.4.1. Specifications of the Adapted Turkish Version of the CCTDI.....	88
Table 4.5.1. Model Fit Statistics across Cultural Groups.....	91
Table 4.6.1. Means and Standard Deviations for the Sub-scales and the Overall Scale across Turkish and American Samples.....	94

LIST OF FIGURES

Figure 2.5.1. Schematic Screening of Cross-cultural Research.....	34
Figure 3.4.1. Multiple Interactive Translation and Back-translation Process.....	54
Figure 3.7.1. Hypothesized Unidimensional Seven-factor Measurement Model of the CCTDI.....	66
Figure 4.4.1. Modified Four-factor Model of the CCTDI with Parameter Estimates for Each Cultural Group Separately.....	87

Chapter 1

INTRODUCTION

1.1 Background of the Study

The phrase ‘critical thinking’ (CT) has been very frequently uttered for the last two decades in educational contexts across United States and Europe as well as Asia. For the recent five years, most of the researchers all over the world engaged themselves in trying to understand what critical thinking actually means, why it is so important, or whether it is directly related to or affecting the education that they provide, how they can embed critical thinking into the educational programs or help their students to engage in critical thinking, whether critical thinking is a product or process laden issue, and most importantly, how they can assess critical thinking or evaluate programs in terms of critical thinking. Besides all of these questions, assessing critical thinking dispositions of pre-service teachers across different cultural groups with a Cross-culturally validated instrument has been one of the most important concerns of a significant number of scholars worldwide (Grosser & Lombard, 2008; Johnson & Reiman, 2007; Lee, 2005; McBride, Xiang, & Wittenburg, 2002; Melnick & Zeichner, 1998; Metzler & Blankenship, 2008; Yeh, 2002). Especially Cross-cultural assessment of the construct of ‘critical thinking’ has been considered to be the most popular subject to be studied (Lee, 2005). For this reason, investigating for a reliable and valid way to assess critical thinking dispositions of pre-service teachers across multiple languages and cultures has

become the major concern of the scholars across the world. Many research instruments have been developed to gain deeper insight into the extent to which the prospective teachers possess the abilities and dispositions of critical thinking. Many scholars following these developments preferred to use existing instruments by adapting them into their language and culture rather than developing new ones. According to some leading researchers in the field, one of the most important reasons that accelerated such Cross-cultural studies is related to understanding whether a proposed conceptualization regarding the construct of critical thinking and, in relation to this, hypothesized assessment model in one language and culture exist in a similar structure in different languages and cultures (Behling & Law, 2000; Hambleton, 2005; Sekaran, 1983; Sireci et al., 2006; Stansfield, 2003). The basic premise behind of their efforts is to seek the extent to which a measurement model designed for one culture be applicable for another one. Of-course, the root of this idea traces back to the curiosity for seeking of a universally accepted criteria for assessing critical thinking of pre-service teachers across the countries, cultures and languages of the world.

It is obvious that considerably incredible efforts have been dedicated to assess and compare critical thinking dispositions of prospective teachers across countries of the world. As it is clear from the noted efforts of scholars, assessing critical thinking dispositions of Pre-service teachers has become an important educational concern worldwide. However, why is it so? Why does it count to assess critical thinking dispositions of Pre-service teachers? Why do researchers frequently utter ‘critical thinking’ together with the terms ‘teacher’ and ‘teacher education’? What is the true meaning of putting such a great emphasis on using these terms together? Answering

these questions, indeed, is necessary for providing a better baseline for the rest of the study.

The analysis of contents of several government reports and critical studies left that the most important reason associated with this great interest given to the assessment of critical thinking dispositions of pre-service teachers is related to the reform strategies carried out in teacher education programs across the U.S and Europe (Benesch, 1993; Brookfield, 1997; Cheong & Loong, 2007; Ellis, 2005; Ennis, 1993; European University Association, 2007a; European University Association, 2007b; Facione, 2006; U.S. Department of Education, 2000). Another critical proposition of these works was that the common element of the reform strategies carried out was fostering critical thinking dispositions of Pre-service teachers. Even Freire (1974), in his famous book *Education for Critical Consciousness*, mentioned the exceptional place of critical thinking as an inevitable and most important part of any reform strategy in teacher development or reform in teacher education. However, it is important at this point to better understand what is actually meant by “reform in education”, why such a reform primarily involves the active inclusion of ‘critical thinking’ into teacher education programs, and what is the intense relationship between the critical thinking reform in education worldwide and the Cross-cultural assessment of critical thinking disposition?

The argument supporting the proposition agrees with the requirements of the era in which we live; therefore, the first question that needs to be answered is that of what does the 21st century demand human being to exist and survive in the system, within which education is *formed* and *reformed* accordingly? While some leading scholars of the related field characterize the current century using phrases such as

“rapid population increase,” “multi-functionality,” “fragmentation and uncertainty,” “non-stability,” and “virtual reality”, they also think that we need to check the intriguing examples that exist in our close environment to know the constructs that determine the human needs for the 21st century (Barton, 2009; Mallik, 2004; Sclove, 2010; Sen, 1997). For instance, to stress the incredible population increase in the world countries, Barton (2009) says that if we took every single job in the U.S. today and shipped it to China, it still would have labour surplus; and in every ten seconds, 60 babies will be born in the U.S., 244 babies will be born in China, 351 babies will be born in India, and 24 babies will be born in Russia. According to the report of the U.S. Department of Labour (2010), a person in today’s world must be multifunctional because the studies show that “today’s learner will have 10 to 14 jobs by age 38” (p. 55). Again, according to the same report, “1 out of 4 workers today is working for a company for whom they have been employed for less than 1 year and more than 1 out of 2 are working for a company for whom they have worked for less than 5 years” (p. 23). Besides, the former secretary of education Richard Riley (as cited in the report of the U.S. Department of Labor, 2010) indicated that the top 10 jobs that were in demand in 2009 did not exist in 2004. This means that educational systems are currently preparing students for the jobs that do not currently exist.

The reflection of technology on social life is incredible as well as exponential. According to a study of Mallik (2004), a lot of unfamiliar terms and technology will be built in this century. As he puts them, the number of text messages sent and received every day exceeds the population of the planet; it is estimated that 1.5 exabyte (1.5×10^{18}) of unique new information will be generated worldwide in 2013; and predictions are that by 2013 a supercomputer will be built that exceeds the

computation capability of the human brain, and by 2023 when first graders will be just 23 years old and beginning their first career, it will only take a \$1000 computer that exceeds the capabilities of human brain (Mallik, 2004). As Mallik continued, to see the relationship between the rationale behind the importance given to the assessment of critical thinking and the current century, more examples and experiences should be generated.

From epistemological point of view, current century prepares a more challenging future for individuals. For instance, the amount of new technical information, as a very well known fact, is doubling every 2 years. For students starting a four year technical or college degree, this means that half of what they learn in their first year of study will be outdated and by their third year of study it is predicted to double every 72 hours by 2014 (Sen, 1997). According to Sen, it is estimated that a week's worth of New-York Times contains more information than a person was likely to come across in a lifetime in the 18th century. Human being is experiencing a dramatic technological, social, and political revolution in the new millennium, which opens up new life-species for all of us. The way we work, communicate with each other, spend time on leisure activities, the way we understand each other, and habits of living have been changed and still been changing.

Change, in this context, is such a term that not only signifies the differentiation of daily habits of living but also connotes with another term for the last three decades, and that is *globalization*. A considerable body of evidence indicate that getting deeper understanding of the 21st century accompanies with unpacking the intense relationship between globalization, as the defining feature of 21st century, and

reform in education (Barton, 2009; Giroux, 2000; Mallik, 2004; Rodrik, 1997; Sclove, 2010; Sen, 1997; Smith, 2006). Yet, very few researchers have so far mentioned the strong bond between globalization, education, and critical thinking, and very little research proposed ways ensuring that globalization may not literally wipe away the educational conventions of societies. At the first glance, one may not see the intense relationship between globalization and essential reform in education in terms of critical thinking. A closer look at these concepts under careful considerations of their influences on one another reveals an important link. In-order to provide a deeper understanding, I find it essential to concentrate on the concept of globalization first.

All around the world, changes happen in the societies, and the common element of these changes is “the international integration of goods, services, and capital” (Rodrik, 1997). Such process of integration is called globalization and those changes are “pressuring societies to alter their practices” (Rodrik, 1997). As a result, people tend to develop new perspectives, new attitudes, and new ways of life accordingly; regardless of the type of change they experience. For this very reason, such economic happenings directly affect the way people organize and live their lives. The question is, what is the direct or indirect relationship, if any, between economics and education? Undeniably teachers are the key players who need to be aware of the changes in their societies (Browne & Freeman, 2000). However, when teachers come across the challenge of globalization, what should their attitudes be? Rodrik (1997) sees globalization as one of the main causes of all social alterations and diversities. Rodrik also contends that continual mobility and change produces diverse societies, with members coming from different countries of the world,

bringing diverse attitudes and diverse ideas with them. Such whirl of diversity equals a diverse culture, in which many different consumption habits are formed or adopted, and different points of views are raised within the same society. Therefore, teachers are bound to deal with diverse populations in this globalizing world, and the teachers' sensitivities as well as their dispositions toward such issues are important (Lau, 1992; Melnick & Zeichner, 1998).

Mark Smith (2006), in *Globalization and the Incorporation of Education*, provides a very well prepared and clear explanation of how globalization seized the arena of education. According to Smith, "forces associated with globalization have conditioned the context in which teachers operate, and profoundly altered people's experiences of education" (p. 2). Smith, first, explains how education has been commodified, and taken over by transnational corporations. Commodification causes people to see education as a good to be consumed so that education becomes a commodity for people, rather than a social or personal need. Most enroll in a university program according to the requirements in the labor market. This causes people to decide on which program to be educated by considering whether they will have a chance in the labor market to have job, and whether the payment of this job will be satisfactory for them. Although Smith does not say that he sees this shift as a negative one, what he implies is that such a change in people's perceptions of education is affecting the practice of education. Second, Smith mentions advancements in educational technologies as another branch of reflections of globalization on education. Though they are not openly visible, the ideas of lifelong learning, individualized learning, distance education, web-based learning all carry the finger-prints of globalization. All of these activities are promoted as part of

individualized programs, and in a de-localized way. As long as people keep on going to their “study tours, fitness centers” and buy their “self-instruction manuals, electronic networks”, they will give way to transnational corporations to keep their hands on education (Smith, 2002). Finally, Giroux (2000) indicates, “it is time to recognize that the true tutors of our children are not school instructors or university professors but filmmakers, advertising executives and pop culture purveyors” (Smith, 2006). This sentence completes the last piece of jigsaw puzzle of globalization in education. In this view, it is not difficult to see how schools are becoming dis-functional to some extent, and how teachers can no longer be as influential as they once were, in helping children experience intellectual transformation and coping with the challenges of the 21st century.

So, there is a new era opened up by the changes being recorded in the current century. Consideration of the current context, as defined above, it is obvious that individuals are required to think and behave on a different level. So, one might still ask who is the ideal person who can cope with the demands and challenges of the new millennium? Tony Wagner (2008), in his famous book, *The global achievement gap: Why even our best schools don't teach the new survival skills our children need--What can we do about it*, provides a considerable effort to outline seven survival skills, and critical thinking is at the top of his list. According to many researchers, critical thinking ability is the key characterological attribute of the 21st century man for several reasons. First, anyone, who wants to catch up with the changes experienced every day, must be open-minded, which is a facet of critical thinking (Lau, 1992; Lee, 2005). Second, understanding, interpreting, evaluating, adapting the new technological advancements require critical thinking abilities

(Barton, 2009; Wagner, 2008). Third, dealing with global problems require global solutions and global solutions can only be generated by having a collaborative effort where individuals have to form new ideas as a result of such patchwork studies, says Bensley (2006) in his study. Fourth, as the countries of the world move towards a technology based economy, having worldwide competition as a result of globalization, the demands of the global economy remind us the importance of personal decision making, reasoning well, and making good judgments. Reed (1998) says, this is such a competition where “employers demand workers who think flexibly and analytically, integrate information from a variety of sources and perspectives, and make profitable decisions efficiently, (p. 2)” and adds that most of the societies of the world today is pluralistic where individuals are required to fair-mindedly evaluate the relevance of various ideas on different problems. Although reasons are not limited with these examples, critical thinking is considered to be the most essential tool for human being to perform different roles successfully in a society which is defined as “fragmented,” “non-stable,” and “uncertain.” In each of these roles, as Perkins (1989) states, human being must

examine the factors impinging on a situation, forecast the outcomes of possible courses of action, evaluate those outcomes and weigh them relative to one another, and try to choose so as to maximize positive outcomes and minimize negative ones. Further, the beliefs we hold, and consequently the inferences we later make and attitudes we later assume, depend in part on our reasoning about the grounds for those beliefs. Accepting beliefs wisely serves the ultimate end of later sound conduct as well as the more immediate end of sound belief itself. (p. 175)

To exist in the system and become a part of it, individuals should go with a serious attitude change from being passive receivers of information to become active interpreter, modifier, and judger of information. So, how should education react or be

designed to help citizens satisfy the needs of 21st century is the question that established the most critical question of the required reform in education all over the world. Therefore, it is now getting clearer to understand the rationale behind embedding critical thinking as a reform strategy in teacher education programs and the need to assess critical thinking dispositions. But a little bit of more effort is required to see the interrelationships and the greater picture. For instance, in this century, John Dewey (1910) asserted that learning to think should be the central purpose of education. According to Scriven, “training in critical thinking should be the primary task of education” (1985, p. 11). Therefore, as a result of the analysis of the bulk of the related literature, unpacking of the important relationship between the requirements of the 21st century, ideal profile of 21st century man, globalization, and education as well as their influences on one another leads to consider a paradigm shift in teacher education in order for education meet the critical thinking need of citizens of societies (Dewey, 1910; Dewey, 1938; Dewey, 2004; Dworkin, 1959; Eklof, 2005; Ernst & Monroe, 2006; Fisher, 2001; Fogarty & McTighe, 1993; Giancarlo & Facione, 2001; Hager & Kaye, 1991; Harrison, 2004; Ikuenobe, 2001; McBride, Xiang, & Wittenburg, 2002; U.S. Department of Education Report, 2000). Lau (1992) also supports this view by indicating, “... students’ performance in thinking will not improve much if the quality of teaching is not much improved. And teaching will not improve much without dramatic improvements in the field of teacher education” (p. 1). Accordingly, it turns out to be a fact that any expected improvement in any society needs the same improvement in teacher education. Therefore, if we seek a paradigm shift in a society and in the way people think, then, we should also seek a parallel change in teacher education programs.

Having said all these things, one can now come to say that it is quite meaningful, highly necessary and quite urgent to make refinements or reforms in teacher education programs in terms of critical thinking and, as the studies show, that is why and what most of the developed countries have been doing for nearly two decades. When our focus of attention; however, has been shifted from other countries of the world to Turkish educational context, an encouraging but insufficient scene is available. Likewise, in a recent report of the Turkish Council of National Education (2009), the renewed goal of higher education, and specifically of teacher education was identified as to train independent, cooperative and proficient educators and teachers, who value critical thinking, believe in lifelong education, have a sense of responsibility for the community at large, and are disposed to internalize contemporary and global values. According to some other leading national reports, reform strategies in Turkey and in Turkish Republic of Northern Cyprus have gained a great momentum to establish greater accountability for critical thinking in teacher education and to embed critical thinking in the curriculum of teacher education programs through accreditation facilities (Turkish Council of National Education, 2009; Turkish Educational Association, 2010; Turkish Kızılay General Education Report, 2011; Turkish Republic of Northern Cyprus Ministry of Education Report, 2005). But, when the time comes to judge our educational context and see where we are within the reforms carried out in the field of teacher education *practically* in comparison to the predefined standards and to the rest of the world, the answer to this question is lacking. As a result of these attempts at undertaking the issue of critical thinking in teacher education programs in Turkish higher education context, an important lack in empirical research and assessment of critical thinking has been

emerged. Existing studies, however, have used only qualitative methods and there was no authorized instrument to assess and compare the critical thinking dispositions of Pre-service teachers across Turkey and other developed countries.

With this lack in empirical assessment, search for a means of ascertaining a reasonably informed opinion about Pre-service teachers' critical thinking dispositions in the Turkish higher education institutions was necessary. The search led to the discovery of the California Critical Thinking Disposition Inventory (CCTDI). This was the only instrument found to be well conceptualized to evaluate teacher education programs in terms of critical thinking dispositions, and translated into many languages such as Arabic, Chinese (Mandarin), Dutch, Farsi, Finnish, French (Canadian), Hebrew, Italian, Japanese, Korean, Portuguese, Spanish (Mexico-Latin America), and Thai. Although the reasons of these nations in selecting this instrument are not limited by those, the CCTDI is unique for two reasons. First it is the only instrument that was designed with the intention of assessing *critical thinking dispositions* of individuals and second it is the only instrument that has currently been translated into more than 12 most spoken languages of the world; which, in turn, allows for assessing the Cross-cultural validity and applicability of the instrument to conduct further Cross-cultural comparative studies. In addition, the development of the CCTDI has been considered to be an important beginning point in the flow of reform in teacher education programs all over the world (U.S. Department of Education, 2000; Facione, Facione, & Giancarlo, 2000; Ikuenobe, 2001; Yoon, Schmidt, & Ilies, 2002). That is why the CCTDI has been translated into more than 12 languages of the world. The CCTDI is also considered to be an inevitable part of the reform process throughout the world countries because scholars

concern themselves to see where they are in terms of the predefined standards for teacher education (Benesch, 1999; Coleman, Rogers, & King, 2002; Facione, 1990; Facione, Facione, & Giancarlo, 1977; Phillips & Bond, 2004). Consideration of the importance of the situation led nations to provide considerable efforts to support such Cross-cultural studies to benefit of their results. Finland and Holland are the two most important countries that started the process of reform in their teacher education programs along with the CCTDI being translated into their targeted languages earlier than other countries in Europe (Facione & Facione, 1992).

Since Turkish educational system is subject to a paradigm shift in teacher education programs, substantial attention has not only been directed to assess critical thinking dispositions of Pre-service teachers in Turkish higher education institutions but also directed to compare the developments achieved in Turkey to other countries with a Cross-culturally validated instrument. Since an authorized instrument measuring critical thinking dispositions and professional judgment in teacher education is not currently available in Turkish language, it becomes essential to set out to adapt the CCTDI from English source language to Turkish target language and assess cross cultural validity across Turkish and American populations.

1.2 Problem to be Investigated

Considering the background about assessing critical thinking dispositions of Pre-service teachers worldwide and in Turkish higher education institutions, and considering the strong theoretical foundation of the CCTDI along with many researchers' suggestion for translating existing instruments rather than developing new ones, carrying out a research on adapting the CCTDI from English source language to Turkish target language has become meaningful (Chapman & Carter,

1979; Waltz, Strickland, & Lenz, 1991; Weeks, Swerissen, & Belfrage, 2007). Especially Waltz et al. (1991) support that the use of existing instruments increases the utilization of assessment tests and decreases the cost of outcome assessment.

The need for a Turkish version of the CCTDI is obvious but when it comes to end up with a multilingual versions of an instrument, the challenge becomes greater. The real problem that needs to be investigated is not simply determining the strategies to translate the instrument and pilot test it for use with the targeted group of audiences. A significant body of evidence of mostly cited scholars in the field of Cross-cultural outlets imply that when attempted to use an existing instrument to measure a phenomenon in an another cultural group and language for which the instrument was not originally developed, the instrument must not only be simply translated into the target language but also be adapted to the cultural group and be assessed on the basis of psychometric properties, Cross-culturally validity and measurement invariance (Ægisdóttir, Gerstein, & Çinarbas, 2007; Chapman & Carter, 1979; Nunnally & Bernstein, 1994; Sireci, Yang, Harter, & Ehrlich, 2006; Stansfield, 2003). For that matter, Hambleton (2005) prefers to use the term *instrument adaptation* rather than *instrument translation* because the term *adaptation* reflects more of the actual process in which many different practical guidelines followed for preparing the instrument for use in a second language and culture. The theoretical framework for this study supports that if the translated Turkish version of the CCTDI is linguistically, functionally, and culturally identical to the original English version of the CCTDI, then, empirical evidence regarding the Cross-cultural equivalency should be obtained. Therefore, the problem turns out to be providing evidence regarding the extent to which the translated Turkish version of the CCTDI

demonstrates linguistic equivalency, structural equivalency, and Cross-cultural applicability.

1.3 Purpose of the Study

The purpose of this study is to investigate empirical evidence supporting the extent to which the adapted Turkish version of the CCTDI is linguistically, functionally, and structurally equivalent to the original English CCTDI.

1.4 Research Questions

The following research questions will be addressed to achieve the purpose provided for the current dissertation study.

1. Given findings regarding the translation and back-translation process, which adaptations are required to end up with a linguistically equivalent Turkish version of the CCTDI?
2. Given findings regarding the necessary statistical analysis, what do both the translated Turkish and the original English versions of the CCTDI demonstrate in terms of their psychometric properties?
3. Given findings regarding the confirmatory factor analysis (CFA), what is the extent to which the data derived from both the Turkish sample and the American sample explained the hypothesized 7-factor measurement model of the CCTDI?
4. Given findings regarding the measurement invariance tests, what is the extent to which the translated Turkish and the original English versions of the CCTDI allow for Cross-cultural mean comparison of the construct?
5. Given findings regarding the adapted CCTDI, what are the critical thinking dispositions of Pre-service teachers across Turkish and American higher education institutions?

1.5 Significance of the Study

The current study is important and necessary to conduct for several reasons. Primarily the necessity to conduct this study is mostly related to the absence of a theoretically supported instrument to measure critical thinking disposition in Turkish higher education context. Specifically, there is no reliable and valid empirical assessment tool in Turkish language, which intends to assess the extend to which pre-service teachers disposed to think critically in the 21st century in which survival mostly depends on how people think. By the end of this dissertation, one of the expected outcomes is to come up with a reliable and valid Turkish version of the CCTDI, which will enable scholars and researchers to assess critical thinking dispositions of pre-service teachers in Turkish higher education institutions. By this way, it will also be possible to conduct situation analyses studies and needs analyses studies in terms of critical thinking in Turkish education context.

The current study does worth to conduct for another significant reason as well. As a part of its outcomes, this dissertation will provide interested researchers with opportunities to conduct Cross-cultural comparative studies regarding critical thinking dispositions with numerous independent variables that can be generated. Such Cross-cultural studies have been widely recognized all over the world and counted valuable since they are considered to be informative with the inferences about cultural differences and similarities regarding the construct being compared. Indeed, by the end of the current research, the results will inform the researchers about the extend to which the CCTDI can be used to allow Cross-cultural mean comparison of the construct being assessed.

One other rationale behind carrying out current investigation can also be outlined as this study to be example for the scientific education milieu in Turkey about the sort of methodology and procedures that need to be utilized when intended to use an existing instrument to measure a phenomenon in Turkish language and culture for which the instrument was not originally developed. Likewise, the search for finding out similar studies in Turkish research context yielded a significant conclusion regarding the issue. The study of Karadağ (2011) reviewed most of the studies from the scale quality as well as translation quality point of views. Accordingly, he found that most of the studies directly translated the tool from source languages to Turkish target language and/or sometimes back-translated it, and then, made instruments available for use in Turkish language only by relying on alpha coefficients and correlation coefficients or time-interval t-test results without signifying the rationale behind. Some studies; however, included confirmatory factor analysis results, which is actually essential, to make reliable inferences regarding the hypothetical measurement model, but this time either the correct procedures for running the test or the appropriate interpretation of the results of analysis were missing. In the case the specified conditions met, most of the studies avoided or did not continue with measurement invariance tests to make inferences regarding the extent to which the translated version explains the culture of the targeted population. On the other hand, most of the studies that translated an inventory or a parametric measurer from a source language to Turkish language are lacking in meeting the necessary standards in test adaptation. The term test adaptation in this context, as defined by Hambleton (2005), includes all the procedures from identification of actual intention of each item to their relevance to content domains, from deciding

who and why will be translating and back-translating to who will be comparing the inner versions, from what standards to be applied to how they will be tested, and so on. The procedures for merging a standardized parametric inventory from one language and culture to another requires extra sensitivity in comparison to developing a new instrument for a particular language and culture. This study will be unique with its methodological nature and be considered as a guideline to conduct a similar study.

Beyond all of the specified reasons, this research will be such an informative baseline that will create some degree of awareness in the field of study. Critical thinking and its assessment has long been a focus of attention. Taking the story from Socrates to John Dewey, most of the leading philosophers paid substantial attention to understanding, practicing, assessing, and developing these traits in individuals. Therefore, every effort in adding something new to the existing literature of our nation equals to increasing our knowledge of the phenomenon of critical thinking. While studies on assessing critical thinking dispositions of pre-service teachers counted very important all over the world, conducting such a study for deepening our understanding of the phenomenon of critical thinking and bringing new insight into the assessment of critical thinking disposition in Turkish higher education context can be considered highly significant.

1.6 Summary

The introduction chapter has established a set for readers in answering their needs for better grasping the rationale behind conducting the current dissertation study. As a baseline, the changing need of human being in relation with the advanced developments, globalization, and changes in technological, social, economic,

political, and psychological conducts has been discussed. In addition to that, the demands of the 21st century, the survival skills that 21st century man needs and necessary practices have been outlined. It is pointed out that the most widespread understanding of solving this problem has associated with an urgent reform in teacher education programs in terms of critical thinking. Following this explanation, the need to assess critical thinking dispositions of pre-service teachers have been raised in order to cross-check the situation in teacher education programs. The rationale with its all necessities and benefits behind adapting California Critical Thinking Disposition Inventory from English source culture and language to Turkish culture and language has been established.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

The intention of literature review chapter is to create an integrated perspective regarding what has already been done and said about the major concerns of the current dissertation study. Since critical thinking takes the most important part of the study, the first sub-section of this chapter includes references from studies and reports for the purpose of bringing a deeper insight into what critical thinking and its components are. The conceptualization of critical thinking and its components also provide readers with opportunities to see the definitional evolvement of critical thinking on a historical base. Following that, studies examining the approaches and techniques to nurture critical thinking dispositions of Pre-service teachers has been placed. This particular section additionally places competing studies in a manner of argument and counter argument analysis. The third sub-section serves readers with research studies intended to develop assessment strategies to assess the critical thinking dispositions and skills of adolescents. In general, the assessment procedures, instruments, and other techniques developed to assess critical thinking have been reviewed with a special focus on their actual purposes of assessment. In specific, their pros and cons have been documented and compared with each other under the light of the related references. The last sub-section, which is dedicated to provide with better understanding of Cross-cultural research phenomenon, is divided into

three sub-headings, namely: (a) approaches to study culture, (b) equivalence, and (c) bias.

2.2 Conceptualization of Critical Thinking and Its Components

The review of the literature on critical thinking showed that scholars since the time of Socrates and Platon have dedicated serious efforts to conceptualize the phenomenon of critical thinking (Ennis, 1991; Hare, 1979; Ryle, 1963; Scheffler, 1966). Such efforts were accelerated with the need of a reform in education across the United States and European countries in the early of 1920^s by enabling critical thinking as a central focus of the reform (Bailin, Case, Coombs, & Daniels, 1999a; Dewey, 1910; Dewey, 1938). Reformists, who were usually the politicians and people in charge of controlling the educational milieu, proposed their own definitions, whereas some other significant group of educators, practitioners and scholars proposed various different definitions and practices for critical thinking (Benesch, 1993; Bailin, Case, Coombs, & Daniels, 1999b). Despite their efforts to come up with a consensus definition, it was most of the time considered to be a challenging task for most of the scholars (Byrne & Jhonstone, 1987). Actually the true challenge in theorizing critical thinking was hidden in the nature of the knowledge of the term. Indeed, some intriguing questions involved: Can critical thinking be directly taught, what actual practices should be utilized to teach critical thinking, how can critical thinking be embedded into the curriculum of the entire educational system, and more specifically, how can critical thinking as an important outcome of education be assessed (Dworkin, 1959; Ennis, 1993). Whilst the discussion between scholars and politicians were severe, researchers among themselves were also divided into many different groups in their utterances regarding

the ideas they put forward, which made the situation even more complicated than ever (Geertsen, 2003). This was especially true when researchers tended to use different labels referring the same phenomenon under investigation. ‘Reflective thinking’, ‘problem solving’, ‘problem screening’, ‘creativity’, ‘creative thinking’, ‘critical thinking’, ‘critical thinking disposition’, ‘critical thinking skill’, ‘higher-order thinking’, ‘lower-order thinking’ and even ‘good thinking’ were some of the examples for those labels that were used interchangeably (Bailin, Case, Coombs, & Daniels, 1999a). The other part of the disagreement was related to discussions involving whether critical thinking is a product laden or a process laden issue (Bailin, Case, Coombs, & Daniels, 1999b), whether critical thinking is a context bound or not (McPeck, 1981), whether critical thinking is a culture specific concept or is there a universally accepted criteria defining the term (Ennis, 1985; Ennis, 1987), whether there is an extent to which an individual be an expert in thinking, and whether critical thinking is a discipline based phenomenon or not (Walker & Finney, 1999). For this reason, conceptualizing critical thinking has long been a jigsaw puzzle for scholars.

The analysis of various paradigms and theories generated for critical thinking revealed that the noted disagreements rooted back to the distinctions between two disciplines, namely philosophy and psychology. While philosophers concentrated on the argumentative nature of critical thinking (Ennis, 1967; Ennis, 1980; Foucault, 1973), psychologists perceived critical thinking as a cognitive process and focussed on more practical definitions (Freire, 1974; Kurfiss, 1988; Pitters & Soden, 2000; Popper, 1935). Moreover, philosophers have based their paradigms on logical reasoning (Ennis, 1980; Foucault, 1973) whereas psychologists most of the time conditioned the context to infer their paradigms as a result of empirical research

(Lewis & Smith, 1993; McIntire & Miller, 2000). Of course, all of these attempts were to conceptualize critical thinking. Most of the educators and practitioners thought that any definition must have hold the fingerprints of both disciplines (Prior, 2000; Roberts, 1998; Said, Adikan, Mekhilef, & Rahim, 2005), and their call was considered to be a turning point for the entire literature. Likewise, Keeley and Browne (1976) proposed the first definition including common conceptions and they considered critical thinking to be:

a reflective skeptical or questioning attitude, a sensitivity to value-or ideology-laden assumptions, as insistence on appropriate supporting grounds before accepting disputable claims, an appreciation of the various criteria applicable to good reasoning and argument (whether general or subject dependent), skill and judgement in the analysis and evaluation of claims and arguments, and a disposition to be self reflective, sensitive to one's own possible biases and assumptions. (p. 46)

According to some of the researchers from both disciplines, Keeley and Browne's definition was lacking since it did not include the most crucial terms such as "interpretation", "evaluation", "analysis", "synthesis", and "inquisitiveness" (Ennis, 1980; Hare, 1979; McPeck, 1981). However, their definition was accepted as a starting point which carried the finger prints of both disciplines.

Although most of the time scholars criticized each others' studies, they were also aware that there was an urgent need to come together, head to head, to get to a consensus definition for critical thinking (Browne & Keeley, 1988; Norris & Ennis, 1989). For this reason, for the last two decades, a considerable effort has been given to conceptualize critical thinking and its components with the inclusion of forty-six experts from various disciplines such as philosophy, psychology, education, sociology, physical sciences, and even medicine. This study was known as Delphi

research supported by the Committee on Pre-College Philosophy of the American Philosophical Association, and conducted by Facione (1990). As a result of this attempt in conceptualizing critical thinking, a report was produced. This report was then recognized as Delphi Report in the related literature. According to this report, the experts involved in this research reached to a consensus definition for critical thinking. This was also known as the most recent conceptualization regarding the phenomenon of critical thinking for the last two decades. According to this report, experts reached the following consensus statement:

We understand that critical thinking to be purposeful, self-regulatory judgement which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgement is based. Critical thinking is essential as tool of inquiry. As such, critical thinking is a liberating force in education and a powerful resource in one's personal and civic life. While not synonymous with good thinking, critical thinking is a pervasive and self-rectifying human phenomenon. The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgements, willing to consider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit. Thus, educating good critical thinkers means working toward this ideal. It combines developing critical thinking skills with nurturing those dispositions which consistently yield useful insights and which are the basis of a rational and democratic society. (Facione, 1990, p. 1)

As Facione (2006) argued, this definition was an operational one which was theoretically strong enough to alter researchers' practices to conduct and carry out empirical research. The other argument proposed by Facione was about the length of the consensus statement. The phenomenon under investigation could not have been defined by any other shorter terms because the phenomenon under investigation had

inevitable connections and contributions to many concepts; therefore, the definition must have involved indications covering those concepts as well.

This definition for critical thinking has been widely accepted and used as an operational definition by many other researchers as well as by this dissertation study. Unpacking this definition reveals that critical thinking is composed of two dimensions, namely *skill* dimension and *disposition* dimension. The California Critical Thinking Disposition Inventory (CCTDI), which attempts at assessing the extent to which a person is disposed to think critically is a result of the Delphi effort and is theoretically based on the experts consensus statement. The rest of the Delphi report explains the development process of the CCTDI with additional definitions for each of its facets regarding the disposition dimension of critical thinking, with special focus on the distinction between skill and disposition dimensions. However, these theoretical and operational frameworks of the CCTDI have been documented in great detail in a separately dedicated section in the methodology chapter of this dissertation.

2.3 Nurturing Critical Thinking of Pre-service Teachers

Critical thinking has been considered to be the central outcome of education, especially of teacher education (Bucy, 2006; Facione, Giancarlo, & Facione, 1995). Likewise, Dewey (1910), in his famous book called *how we think*, asserted that the central purpose of education should be nurturing critical thinking. For this reason, researchers have conducted many experimental studies in search of ways to nurture critical thinking dispositions of Pre-service teachers. Yet the results of their studies are inconclusive regarding approaches or models to nurture critical thinking of teacher candidates (Grauerholz & Holtrop, 2003; Gruber & Boreen, 2003). When the

discussion is on *how* rather than *what*, it is believed that the Pre-service teachers' critical thinking dispositions cannot be nurtured through classroom activities, assignments, or other curriculum oriented routines (Gruber & Boreen, 2003).

According to the pioneers of teacher education, the current philosophy of education, educational administration and management policies, curriculum and instructional methodologies being employed were all insufficient in creating a change in the attitudes of students from being passive receivers of information to become an active interpreter of knowledge and to become effective formers of ideas (Tabachnick & Zeichner, 1991; Liston & Zeichner, 1991; Zeichner, 2009). On the other hand, the effect of various instructional materials, teaching models, and programs have been reviewed by researchers to conclude on the improvement of critical thinking dispositions of Pre-service teachers. McMillan (1987), who was widely quoted by many research studies in the field of critical pedagogy, put the most famous 27 of these studies under objective and investigated the factors enhancing critical thinking dispositions of individuals. The result of his revisions of many studies revealed that there was no such a technique or model of teaching that can successfully nurture critical thinking dispositions of individuals.

Subsequent developments in the chronological order of the literature showed that because the researchers could not get their expected results from the materials and various techniques they experimented, the discussion moved from *trial and error* method to a more phenomenological and more theory based solutions (Suliman, 2006; Tapper, 2004; Wessel & Williams, 2004). This approach then created new divisions of ideas regarding where the problem is and how the problem should be solved. A group of researchers believed that critical thinking should be directly

taught to have an explicit effect on the cognition of individuals (Plath, English, Connors, & Beveridge, 1999). The rationale behind of their indications was that the curriculum of any given teacher education program was highly busy with theory driven courses and students were not provided with opportunities to practice thinking or forming new ideas. Therefore, critical thinking dispositions could not be nurtured within such a strictly oriented curriculum where enough time was not allocated to achieve so. Whereas another group of researchers just thought the opposite. They thought that critical thinking cannot be taught and should not be taught as a separate course in the curriculum (Sellnow & Ahlfeldt, 2005; Shaughnessy, Zechmeister, & Zechmeister, 2003). Their rationale was that critical thinking was an inevitable part of every single detail of the entire curriculum and the facets surrounding it should be available at every small segment of the entire curriculum, not just as an independently organized part of the curriculum. Contemporary concerns; however, regarding the issue disclosed that neither the former approach nor the later one were triumphant in their aims at leading teacher candidates to acquire the necessary critical thinking dispositions. Most recent developments in the field showed that both the problems and solutions are somewhere else. For instance, Zeichner (2009) pointed out that acquiring positive dispositions regarding the use of critical thinking was not directly about the materials, approaches, or any other external or internal forces but rather highly related to the existence of scientific and academic connectedness. In other words, the surrounding culture and synergic milieu, according to Zeichner, should contain the nuances of the construct so that individuals will then create the critical thinking culture. Therefore, the phenomenon of critical thinking is somehow a result of a natural evolvement. According to that explanation,

there is no way to nurture critical thinking dispositions with external forces or by changing some independent variables with the intention to have effects on the construct. Scholars, who are in support of this view, add that professional judgement and critical thinking, which are two important attributes of an effective teacher, can only be achieved through establishing a set where individuals can come to question the world around them and answer independently without any structured curriculum (Miri, Chaim, & Uri, 2007; Walker & Finney, 1999). With existing curriculum structures, any external attempt will always be inconclusive.

The view of Zeichner has been widely accepted by many authorities all over the world and contemporary efforts have been dedicated to alter their changes in the proposed directions. On the basis of his suppositions, countries such as Finland, Singapore, Hong Kong, and Canada were the ones who changed their learning and teaching culture from behaviorist structure to more flexible and need oriented innovative learning culture (Walker & Finney, 1999). For instance, Finland reduced the theoretical course hours from twelve to three hours and increased the practical observation and outside class activities from one hour a day to five hours a day (Wagner, 2008; Walker & Finney, 1999). The projects generally included observations of nature and keeping track of how the earth as a part of the universe works. They believed that this way of treating education will give birth to a more thinking and creative culture where critical thinking will automatically be available as a result of the actual and natural process. Likewise, researchers conducted longitudinal studies most of which aimed at collecting data and reporting empirical evidence regarding the issue. As Zeichner (2009) concluded critical thinking cannot be taught, it does already exist. The only thing required is to give individuals the true

opportunities to practice and develop their dispositions critical thinking. As a fruitful area of research, the findings regarding the new approach are expected to be shared with the existing educational milieu and literature within the next ten years of time.

2.4 Assessing Critical Thinking Skills and Dispositions

Building standards for what and how should be taught at different levels of any educational system along with an inventory to assess whether or not the predefined standards have been achieved have always been the major concern of researches in the field of educational psychology and critical pedagogy (Shepelak, Jackson, & Moore, 1992; Thomas, 1993; Paul & Nosich, 1992). Setting standards for curriculum and instruction has always been difficult and getting even more difficult as we all experience rapid social, economic, political, and technological advancements and changes. However, assessment still possesses the number one priority in developing programs and curriculum for the shake of nurturing critical thinking dispositions. The reason behind of such an importance given to assessment part of the story, in fact, is very simple. If you cannot assess critical thinking dispositions in a valid and a reliable way, you cannot make inferences regarding the effectiveness of any program developed to nurture the critical thinking dispositions of individuals (Ennis, 1991; McPeck, 1981). This is truly valid for any construct other than critical thinking or any facet of any given educational program. For this reason, considerable number of efforts were dedicated to develop assessment tools to assess critical thinking with and without theoretical supports behind (Murphy, Conoley, & Impara, 1994). Review of the related literature led to combine these efforts under three main categories: (a) standardized assessment tools developed with commercial purpose, (b) standardized assessment tools developed without

commercial purpose, (c) training students to assess or reflect on their own thinking. Within the mainstream of this section, these efforts will respectively be taken under close objection.

Some standardized assessment tools developed with a commercial purpose were evaluated in the national report of the U.S. Department of Education (2000) and namely included: the California Critical Thinking Skills Test (CCTST) developed in 1993 and published by California Academic Press; the California Critical Thinking Disposition Inventory (CCTDI) developed in 1992 and published by California Academic Press; the Watson-Glaser Critical Thinking Appraisal (WGCTA) developed in 1994 and published by American College Testing Program. Except for the CCTDI, which attempts to assess the disposition dimension of critical thinking with a six choice Likert type dichotomous scale, the CCTST and the WGCTA aim at assessing the skills dimension of critical thinking by relying on multiple choice response format. Three of the noted assessment tools commonly cover important facets of the critical thinking construct, namely: ‘analysis’, ‘inference’, ‘interpretation’, ‘evaluation’, ‘reasoning’ and ‘synthesizing’, whereas the CCTDI uniquely covers ‘truthseeking’, ‘open-mindedness’, ‘maturity of judgement’. Despite the fact that they were frequently used for assessment of critical thinking worldwide, none of them indicated that they covered all of the facets of critical thinking (Facione, Facione, & Giancarlo, 2000; U.S. Department of Education, 2000). The reliability and the validity of these instruments were tested by their developers and the developers reported strong psychometric properties (Facione, Giancarlo, & Facione, 1995; Murphy, Conoley, & Impara, 1994).

The other group of standardized assessment tools developed without a commercial purpose were also evaluated in the national report of U.S. Department of Education (2000) and namely some examples included: Academic Profile (AP) developed in 1989 and published by Educational Testing Service; College Assessment of Academic Proficiency (CAAP) developed in 1988 and published by American College Testing Program; College Basic Academic Subject Examination (CBASE) developed in 1990 and published by The Riverside Publishing Company; College Level Academic Skills Test (CLAST) developed in 1984 published by Florida State Department of Education; College Outcome Measures Program Objective Test (COMPOT) developed in 1976 and published by American College Testing Program; Critical Thinking Assessment Battery (CTAB) developed in 1997 and published by American College Testing Program; and all of these tests designed to measure critical thinking aspect in common and drew their judgments by relying on multiple choice item format.

Keeley and Browne (1986) argued that multiple choice response format is not a valid choice to draw inferences regarding critical thinking because responses are limited to five or six choices; therefore, test takers cannot use their creativity or their own criteria to form their own ideas regarding the facets of critical thinking. Besides that, some researchers suggested to train students to reflect their own thinking using student generated scales or by essay writings (Browne & Keeley, 1988; Keeley & Browne, 1986; Norris & Ennis, 1989; Paul & Nosich, 1992). For this very reason, and as an example to the third category, Ennis and Weir (1985) developed the Ennis Weir Critical Thinking Essay Test (EWCTET), which was especially designed to let students read essays on real life basis. These essays intentionally included real life

situations and many thinking errors in it, which was where students were expected to form their own critical responses. In this type of assessment, it was believed that test-takers were provided with much more opportunities to display their critical thinking abilities and dispositions (Ennis & Weir, 1985). Although the EWCTET was considered to be a more adequate way of assessing critical thinking skills and dispositions, to read the responses and making inferences out of the responses were highly time consuming and could not be applied to big populations, but rather preferred to be used with small samples, especially for formative or summative evaluations in a course based environments (Murphy, Conoley, & Impara, 1994). In addition to all, teaching students to evaluate their own thinking based on the determined standards is still a fruitful area of research and researchers pay attention to the necessity for developing strategies to help students assess and reflect their own thinking with a more practical way (Douglas, 2000; Dumka, Stoerzinger, Jackson, & Roosa, 1996).

Recent developments and attention given to the phenomenon of critical thinking summarized that the many ways generated for assessing CT and efforts at developing better strategies will continue because there is no best way to define and assess CT up-until present (Gabennesch, 2006; Gadzella, Hogan, Masten, Stacks, Stephens, & Zascavage, 2001). However, there is another group of researcher who raised another discussion point. They believed that critical thinking or any other form of thinking, whatever label given to it, cannot be assessed with quantitative efforts since they asserted that critical thinking or thinking itself is not something countable (Bailin, Case, Coombs, & Daniels, 1999b; Benesch, 1993; Bensley, 2006). The essence of this argument conveys the nuance that one cannot count how much a

person thinks. However, the counter argument replied that the intention was not to count how much a person thinks, rather, was to count the frequency of observable behavior associated with the predefined criteria of critical thinking (Brookfield, 1997; Browne & Freeman, 2000; Ennis, 1993; Facione, 1990).

The final remark to make about the issue of assessing critical thinking skills and dispositions is that there are many research instruments developed to assess the construct of CT. Each of them has pros and cons associated with their purposes of development. Most of them possess strong psychometric qualities and can be used for program assessment. But, the CCTDI, as the bulk of the literature asserted, is such a unique toll for three important reasons. First, the CCTDI intended to assess the seven-facets of critical thinking disposition not the skill dimension (Facione & Facione, 1992; Paul & Nosich, 1992). This puts the CCTDI out of the ‘countable construct’ or ‘un-countable construct’ discussion. Second, the CCTDI, with a strong theoretical framework and highly positive psychometric properties, serves as a consistent tool (U.S. Department of Education, National Center for Education Statistics, 2000). Third, the CCTDI was first developed in English language by Facione and Facione in 1990 on a basis of a strong theoretical framework and a two years of Delphi study supported by the American Philosophical Association. A year after the study touched down to end the development program, researchers from all over the world attempted to translate the CCTDI from English source language to their own target languages to assess the extend to which students enrolled in different higher education programs, especially in teacher education programs, posses critical thinking dispositions. This effort has been continued for the last two decades and been accelerated for the last ten years. It is a toll that establishes a basis for Cross-

cultural comparative studies with its more than 12 translated language versions (Jones, 2007).

2.5 Cross-cultural Research

With globalization, there is a great interest in Cross-cultural research and international studies in the fields of education and psychology (Sen, 2007; Shaughnessy, Zechmeister, & Zechmeister, 2003; van de Vijver & Leung, 1997). But a greater interest has been found on validation studies (Hsueh, Philips, Cheng, & Picot, 2005; Norris, 1989; Sireci, Yang, Harter, & Ehrlich, 2006). Scholars tended to explore whether developed paradigms or hypothesized measurement models exist across different cultural groups (Hambleton, 2005; Sireci, Yang, Harter, & Ehrlich, 2006). This particular concern of scholars raised discussions of three important

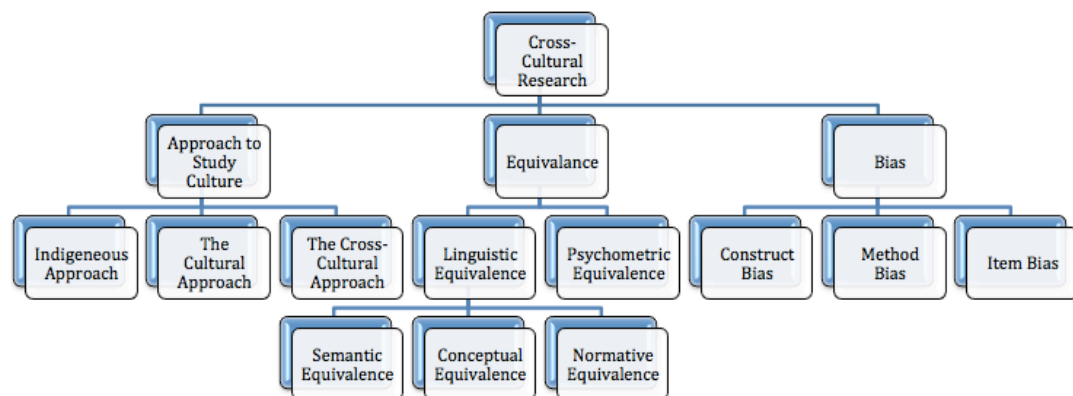


Figure 2.5.1. Schematic Screening of Cross-cultural Research

concepts such as “*approaches to study culture*”, “*equivalence*”, and “*bias*” in Cross-cultural research. For that matter, each of these three concepts will respectively be documented under the light of the related literature. It should; however, be noted at this point that the categorization exposed in this section is not directly available in the outlet of the literature, but rather is a result of the current meta-analysis of the

related literature. The flow of discussion for this section and the schematic screening of Cross-cultural research can also be followed by Figure 2.5.1.

2.5.1 Approaches to Studying Culture

A group of researchers in the field of developmental psychology and Cross-cultural research documented that in order to say anything about the approaches to studying culture, it is essential to define the term ‘*culture*’ (Giroux, 2000; Grosser & Lombard, 2008; Perkins, 1989). For culture, most of the scholars accepted and used the definition provided by Ponterotto, Casas, Suzuki, and Alexander (1995). According to these researchers, culture is a “learned system of meaning and behavior passed from one generation to the next” (p. 4). That is to say, culture has great influence on how people in a particular place attach meanings to the world around them and behave in certain ways. In relation to that, the approaches studying culture and variables associated with culture involved *indigenous approach*, *the cultural approach*, and *the Cross-cultural approach* (see Figure 2.5.1.). The *indigenous approach* enabled researchers to focus on conceptual meanings within a specific cultural context (Adamopolous & Lonner, 2001; Berry, 1969). Studies with indigenous approach generally involved concerns regarding the variations of concepts across demographics of that particular cultural group (i.e., what does *education* mean in *Turkish* culture?), and generalization typically based on the studied culture by excluding outside generalizations. The *cultural approach*, in contrary, shifts the focus of attention from studying their culture to studying other cultures. Although experimental methods were rarely used, the actual methodology involved ethnography and anthropology (Berry, 1989; Wolcott, 2001). In this sort of study of culture, which is similar to *indigenous* approach, researchers aim at

deepening their insight into the meanings of constructs for other cultural groups (Silverman, 2001). The typical action of research involves immersion into the targeted cultural group and collecting data through interviews and observations to make inferences regarding the construct under investigation (i.e., How does Turkish adolescents perceive education?). Here, the researcher has always been an outsider to the targeted cultural group (Silverman, 2001; Wolcott, 2001). In *Cross-cultural approach*, however; researchers study more than one culture and the aim is usually to provide empirical evidence regarding the existence of the construct across different cultures (Brislin, 1976; Brislin 1983; Brislin, Lonner, & Thorndike, 1973). In this type of approach, researchers begin their investigations with an assumption that the construct being studied exist across all cultures considered for the study. Although there has never been a one to one fit between cultures for the constructs being studied, the typical expectation was to provide evidence regarding the extent to which the construct existed across cultures (Ægisdóttir, Gerstein, & Çinarbas, 2007). One other remark that the related literature asserted about the approaches to studying culture was that the indigenous and the cultural approaches were for studying “*emic constructs*”, constructs that are unique to the studied culture whereas the Cross-cultural approach was used to study “*etic constructs*”, constructs that are common for multiple cultures (Berry, 1989; Brislin, 1976). According to Berry, the major goal of the Cross-cultural approach is to seek similarities and differences of a construct across cultural groups (i.e., What are the similarities and the differences between Turkish and American adolescents in terms of their perceptions of education?) and to investigate the comparability of constructs across cultural groups (i.e., What is the

extend to which the construct critical thinking be assessed across Turkish and American cultural groups?).

2.5.2 Equivalence

The literature mentions two types of equivalence in Cross-cultural assessment research. They are namely (a) *linguistic equivalence* and (b) *psychometric equivalence* (see Figure 2.5.1.). Both of them were considered to be the backbones of Cross-cultural assessment research. Moreover, many scholars suggested that any research investigating the existence of a hypothesized measurement model or construct across cultural groups should consider both types of equivalence (Hambleton & de Jong, 2003). For this reason, this section of the literature review has been dedicated to document the most recent developments and indications of scholars regarding these concepts respectively.

According to most of the leading studies, the linguistic equivalence involved *semantic equivalence*, *conceptual equivalence*, and *normative equivalence* (see Figure 2.5.1.). The first subtype of linguistic equivalence, which is semantic equivalence, is concerned with the words and phrases that are identical or similar in their meanings across cultural groups (Hambleton & de Jong, 2003). Supposedly, if a word or a phrase has a different meaning in another language, then, one cannot talk about linguistic equivalency in terms of semantic meaning of that particular word or phrase in that of targeted language. Behling and Law (2000) argued that achieving semantic equivalency or solving semantic problems may be seen easy at the first sight but they consider this process as a challenging one. The best example for the challenge of ensuring semantic equivalence have been found in the studies of Saito, Nomuro, Noguchi, and Tezuca (1996). Saito et al., showed a considerable effort in

achieving a Japanese version of the “Family Environment Scale but translators encountered with serious problems when they attempted to come up with a semantically equivalent version. The bulk of the literature, for solving semantic problems, suggested to use translation-back translation technique developed by Brislin (1970) to reduce the chance of violating semantic equivalency. Many authors’ judgments of the various translation techniques are in support of translation-back translation technique in comparison to simple-direct translation technique, modified-translation technique, parallel-blind technique, random probe technique, and ultimate test technique (Herrera, DelCampo, & Ames, 1993; Wang, Lee, & Fetzer, 2006). Scholars based their judgments about this issue by analyzing these techniques according to their relevance to the four criteria of usefulness, namely: informativeness, source language transparency, security and practicality. According to Behling and Law (2000), the translation and back translation technique was found to be the one that met these criteria at the highest degree.

The second subtype of linguistic equivalence, which is conceptual equivalence, is consistent with the way the construct operationalized across cultural groups (Brislin, 1986). For example, the idea of self-concept refers to individualized persons in western cultures that are identified as individualistic cultures whereas the same concept refers to a person’s family in collectivistic eastern cultures (Berry, 1989). Therefore, any scale including statements about the self concept are interpreted differently by different members of different cultural groups. Solving conceptual problems in the process of Cross-cultural test adaptation; therefore, include consideration of the actual use of the phrases rather than the usage of the phrases across cultural groups. This way of undertaking the issue enables the

researchers to achieve more equivalent cultural forms of the same instrument. This is truly accurate with reconstructing the item for different cultures but still validly assessing the same construct under investigation. In his famous study, Brislin (1970) indicated that the biggest challenge in assuring of the conceptual equivalence across cultural versions is to find the identical phrases for both cultural groups and the construct being measured for those cultures. For most of the scholars (Adamopoulos & Lonner, 2001; American Psychological Association, 2003; Asner-Self & Marotta, 2005; Betz, 2005; Brislin, 1976), translation-back translation method developed by Brislin (1970) should be considered together with a multilevel translation process, which is discussed in depth in the methodology chapter of this dissertation, for ensuring the conceptual linguistic equivalency between the cultural versions of the instruments.

The third subtype of linguistic equivalency, which is known as normative equivalence, is more concerned with the norms of a particular culture or society for which the construct is attempted to be measured (Hui & Triandis, 1983). It is a well known fact that every society has its own set of norms and conventions and that these conventions are quite influential in determining the way individuals operate and behave in that particular society. The degree to which members of a particular culture or society are open to discuss some topics may be different from another cultural group of people (Greenfield, 1997). Especially the manner in expressing the ideas might be different (Hui & Triandis, 1983). In this sort of situations, Cross-cultural research is so sensitive to keep track of those normative differences or similarities to end up with a normatively equivalent versions of an instrument (Asner-Self & Marotta, 2005). Among the constructs accepted to cause normative non-equivalence

were *sexual topics, conformity, nonconformity, assertiveness, personal matters, family oriented matters, political matters, positional matters, and religious matters* (Asner-Self & Marotta, 2005; Behling & Law, 2000; Greenfield, 1997; Hui & Triandis, 1983). Supposedly, respondents of a culture might not be willing to respond the instrument if the manner of the questions or the construct itself does not fit the norms of his/her cultural conventions. This situation creates normative non-equivalency. To sort out such normative non-equivalency between cultural groups, scholars suggested to rewrite the items by modifying the delivery manner without distracting the actual intention of the item (Allalouf, Hambleton, & Sireci, 1999). Researchers also argued to navigate these modifications with experts and translators from both target and source cultures (Behling & Law, 2000; Brislin, 1970).

Remembering from the conceptualization of the types of equivalency, the second type was the psychometric equivalency. Psychometric equivalency refers to the similarities between the psychometric properties of the cultural versions of any research instrument (Dimitrov, 2010; Hambleton, 2005; Nunnally & Bernstein, 1994). Psychometric properties include all types of reliability and validity issues in measurement and evaluation. However, for Cross-cultural psychometric comparisons, only the suggested reliability and validity procedures were considered (Cheung & Rensvold, 2000; Dimitrov, 2010). Though which tests need to be run alternatively depend on the circumstances under which the investigation was carried out (Nunnally & Bernstein, 1994). However, literature suggested to seek equivalency between internal consistency reliabilities, test-retest reliabilities, content validities and construct validities of cultural versions of any given inventory or instrument (Ægisdóttir, Gerstein, & Çinarbas, 2007; American Educational Research

Association, American Psychological Association, & National Council on Measurement in Education, 1999; Dimitrov, 2010; Hambleton, 2005; Nunnally & Bernstein, 1994). The psychometric theory of Nunnally and Bernstein argued to further comparison to advance and include Cross-cultural factorial validity and measurement invariance across cultural groups. Statistical analysis of invariance was also accepted to provide empirical evidence regarding the degree to which the results of different cultural versions of any given test could be compared (Anastasi & Urbina, 1997; Hambleton, 2001).

2.5.3 Bias

Bias is one of the most important concepts that took its place in the discussion of Cross-cultural research. Scholars studying bias pointed out that there is a negative relationship between bias and equivalency and that they are two opposite terms (Johnson, 1998; Kim, 2001; Pedersen, 2003; Triandis, 1976; Triandis 1994). Bias as explained causes non-equivalency between the cultural versions of any test and limits the Cross-cultural comparability of test scores (Triandis, 1976). However, to better understand the concept of bias in cross-cultural research, literature mentions three sources of bias, namely: *construct bias*, *method bias* and *item bias*. As Ægisdóttir, Gerstein, and Çinarbas (2007) assert “a construct bias occurs when the construct measured as a whole (e.g., intelligence) is not identical across cultural groups” (p. 194). This type of bias, in general, refers to the inappropriateness of the definitions provided for the domains of the constructs being measured for each culture. If there is a seriously different conception or coverage of the same construct from two different cultures, then, most likely there will be construct non-equivalency across defined cultural groups (van de Vijver & Leung, 1997). In addition, van de

Vijver (1998) argued that although construct equivalency was established and verified by similar factorial structure across cultural versions of an instrument, which meant there was no construct bias, bias might still exist. van de Vijver called this type of bias as method bias and further indicated that method bias distorts comparability of results as a consequence of measurement-unit or scalar non-equivalence. The root of any method bias was associated with the characteristics of instruments or the way it has been administered to the targeted population (van de Vijver, 2001). Item bias; however, may result from weak relatedness of an item to the content domain, inadequate translation, poor item formulation, and complex wording. Unlike construct bias, item bias does not usually cause overall non-equivalency between cultural versions of the instrument to some extent, yet negatively influences the inter-item reliability (van de Vijver & Poortinga, 1997). As the final remark about the issue of Cross-cultural research, variations in this particular area of study exist for the terminology being utilized, yet the rationales behind each framework rooted back to the same main sources.

Chapter 3

METHODOLOGY

3.1 Introduction

This chapter of this dissertation intends to explain the sort of methodology utilized to tackle with the research problem and answer the research questions with relative theoretical framework underlined the study. For that matter, the first section was dedicated to explain the general research design and theory supported the current dissertation study. Following that, a special effort was given to document the theoretical and operational construction of the CCTDI, which was necessary to provide readers with detailed information of the instrument's developmental process and its' measurement characteristics for comprehensible inferences to be made out of the rest of the sections. The third section housed the translation and back-translation process as one of the most important parts of the methodological flow of this study. In that particular section, every small single detail regarding how the translation of the CCTDI was held was explained in advance. This chapter also included a separate section that involved all the procedures regarding data sampling and data collection by specifying the demographic distributions of the dual validation sample. The following three sections were respectively dedicated to explain the procedures for psychometric analysis of the linguistic versions of the CCTDI, procedures for assessing the cross-cultural factorial validity of the CCTDI, and procedures for assessing the Cross-cultural measurement invariance of the CCTDI. Finally the

current chapter ends with a summary for providing the readers with an overall picture of the entire methodology being employed.

3.2 Research Design

The current study utilized a cross-sectional, descriptive empirical research design, supported by the Cross-cultural measurement theory and psychometric theory (Hambleton, 2005; Nunnally & Bernstein, 1994; Sireci, Yang, Harter, & Ehrlich, 2006) to adapt and Cross-culturally validate the CCTDI. The main premise behind employing Cross-cultural psychometric methodology is to provide evidence regarding whether the results that will be obtained from both source and target language versions of the inventory are because of the errors in translation or true differences in the participants or the variables being measured (Hambleton, 2005). Chapman and Charter (1979) stated that Cross-cultural equivalency could be investigated by examining the measurement invariance across cultural groups.

A considerable evidence has accumulated to provide careful directions for adapting educational and psychological research inventories. For instance, The American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME), in *Standards for Educational and Psychological Testing* (1985), stated the standards for selecting, developing, adapting, and using educational and psychological inventories. Two important basic standards those reports shared and applied to this research were (1) Standard 13.4. “When a test is adapted from one language and culture to another, its reliability and validity for the uses intended in the linguistic groups to be tested should be established” (p. 18); (2) Standard 13.6. “When it is intended that the two versions of dual-language tests be comparable, evidence for test comparability

should be reported” (p. 18). According to the underlying theoretical framework and standards provided in those sources, the process of adaptation involved the followings respectively: (a) translation and back-translation to ensure of linguistic equivalency, (b) assessment of psychometric properties to elicit basic characteristics of an inventory, (c) assessment of Cross-cultural factorial validity to ensure the extent to which the hypothesized measurement model (M_h) exist across cultural groups (Cross-cultural Equivalency), and (d) assessment of measurement invariance of an inventory to ensure of the extent to which the mean scores obtained from both linguistic versions of an inventory are comparable.

Those standards, which are considered to be establishing procedural flow of the theoretical framework for the current dissertation study, are in line with significant group of researchers who are also considered to be the pioneers in the field of Cross-cultural test adaptation (Ægisdóttir, Gerstein, & Çinarbas, 2007; Chapman & Carter, 1979; Dumka et al., 1996; Hambleton, 2005; Nunnally & Bernstein, 1994; Sireci, Yang, Harter, & Ehrlich, 2006; Stansfield, 2003).

3.3 Theoretical and Operational Construction of the CCTDI

Any psychological inventory designed for assessing a phenomenon for educational and research purposes needs to be based on a well-structured and well-conceptualized theoretical framework. For this reason, it is not always easy to end up with a well-conceptualized assessment tool. However, for the last two decades, a considerable effort has been given to conceptualize critical thinking and its components. A cross-disciplinary Delphi study, which was supported by the Committee on Pre-College Philosophy of the American Philosophical Association and included 46 international critical thinking experts, was conducted by Facione

(1990), lasted for two years, yielded a consensus definition of critical thinking. According to the Delphi report, critical thinking (CT) is composed of cognitive skills dimension and affective dispositions dimension, thus, involves *willing* and *able* to use one's cognitive powers of analysis, interpretation, inference, evaluation, explanation, and self-monitoring meta-cognition to make purposeful judgments about what to believe or what to do in a given context (Ennis, 1993; Dewey, 1910; Facione, 1990; Facione, Giancarlo, & Facione, 1995).

If we are to unpack this definition, we understand that in order for a person to make purposeful judgments about what to believe or what to do in a given context, he/she not only needs to have cognitive skills such as 'interpretation', 'analysis', 'evaluation', 'inference', 'explanation', and 'self-regulation' but also needs to be positively disposed to use these skills (Dewey, 1910; Facione, 1990; Lewin, 1935).

Being disposed refers to the affective dispositional dimension of critical thinking. An example for the distinction between a habit and skills exist in different cases of daily life. A person, for instance, who is habituated to healthful living, is more likely to engage in sport activities, eat healthy foods, follow magazines about health, and avoid risky activities (e.g. smoking, drugs, stress...). Another person, on the other hand, might hold the beliefs and skills to engage in the same practice, which is called the healthful life style, but not habitually do so. In that case, we would say that this person is not positively disposed to engage in such practices. The same is valid for thinking. As Facione, Facione, and Giancarlo (1997) explain, people may have the skill to think well or deal with a given problem, and yet, unless some external force demands it, they may not apply their skills to solve the problem. In conceptualizing critical thinking disposition, this example shows that such people do not have strong disposition toward critical thinking and they are not internally

motivated to use their cognitive skills to make purposeful judgments about what to believe or do in a given situation.

The review of the Delphi report reveals that the definitions provided for critical thinking and critical thinking disposition somewhat traces back to the documentations of John Dewey, Karl Popper, and Paulo Freire. For instance, Dewey describes the dispositional aspect of thinking as “personal attributes” (Dewey, 1910). According to Popper (1935) and Freire (1974), critical thinking attributes should primarily be considered as a reform strategy in education instead of critical thinking skills. Facione, Giancarlo, and Facione (1995) further suggest that there is a “characterological profile, a constellation of attitudes, a set of intellectual virtues, and a group of habits of mind which we refer to as the overall disposition to think critically” (p. 2). In the Delphi study, these intellectual virtues and habits of mind have been characterized as ‘truth-seeking’, ‘open-mindedness’, ‘analyticity’, ‘systematicity’, ‘inquisitiveness’, ‘critical thinking self-confidence’, and ‘maturity of judgment’. These virtues are considered as the *characteristics* of an ‘ideal critical thinker’. Further effort, indeed, in defining ideal critical thinker has ended up with the following definition:

The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are precise as the subject and the circumstances of inquiry permit. (Facione, 1990, p. 3)

The CCTDI, which is the end product of the Delphi effort, aims at assessing disposition dimension of CT. According to the 46 critical thinking experts in the Delphi study, the CCTDI represents a high degree of fit between the current conceptualization and measurement development of critical thinking dispositions.

The seven affective dispositions that the CCTDI attempts to assess are shortly defined as follows:

1. Truth-seeking: is to “seek the truth, courageous about asking questions, and honest and objective about pursuing inquiry even if the findings do not support one’s interests or one’s preconceived opinions”.
2. Open-Mindedness: is to be “open-minded and tolerant of divergent views with sensitivity to the possibility of one’s own bias”.
3. Analyticity: is to be “alert to potentially problematic situations, anticipating possible results or consequences, and prizing the application of reason and the use of evidence even if the problem at hand turns out to be challenging or difficult”.
4. Systematicity: is to be “organized, orderly, focused, and diligent inquiry in inquiry”
5. CT Self-Confidence: is referred to “the level of trust one places in one’s own reasoning processes”.
6. Inquisitiveness: is to have “intellectual curiosity by means of valuing being well informed and learning even if the immediate payoff is not directly evident”.
7. Maturity of Judgment: is to make “reflective judgments based on cognitive maturity and epistemic development” (Facione & Facione, 1992, p, 11-12).

Assessing the disposition dimension of CT has gained more importance than assessing the cognitive skills dimension. John Dewey, in *How We Think*, expresses, “If we were compelled to make a choice between these personal attributes and knowledge about the principles of logical reasoning together with some degree of

technical skill in manipulating special logical processes, we should decide for the former” (1933, p.34). Moreover, the motivational theory of Kurt Lewin presents the theoretical framework for the assumption that the disposition to value and employ CT would impel an individual to lead mastery over CT skills, being motivated to close the gap between what is valued and what is attained (Lewin, 1935).

As explained above, significant effort has been given to conceptualize critical thinking and to establish the theoretical foundation of the CCTDI. However, the development of the CCTDI further continued with several other efforts. Development continued with generating measurement items from each of these 7 content domains (see Delphi Report at Insight Assessment) that represent 7 dispositional aspects of critical thinking, which, in turn, established the unidimensional assessment model after the necessary pilot tests and factor analyses carried out within the mainstream of the Delphi effort.

When operationally evaluated, the CCTDI is composed of 75 items rated on a 6 point, forced choice scale (1 = totally disagree, 2 = disagree, 3 = partially disagree, 4 = partially agree, 5 = agree, 6 = totally agree) and intends to measure 7 dimensions of critical thinking dispositions with 7 sub-scales. The 6 point dichotomous forced choice scale intended to group respondents into two main categories such as agree and disagree, and intended to measure the extend of agreement or disagreement within each category.

The Delphi study reported alphas for the sub-scales of the CCTDI: (1) Truth-seeking (12 items, $\alpha = .72$); (2) Open Mindedness (12 items, $\alpha = .73$); (3) Analyticity (11 items, $\alpha = .72$); (4) Systematicity (11 items, $\alpha = .74$); (5) Critical Thinking Self-Confidence (9 items, $\alpha = .78$); (6) Inquisitiveness (10 items, $\alpha = .80$); (7) Maturity of Judgment (10 items, $\alpha = .75$); and overall scale (75 items, $\alpha = .90$).

The scale scores of the CCTDI range between 10 and 60 and interpreted as follows: Scale scores in the 10 to 29 range refer to the low disposition; scores in the 30 to 39 range refer to the ambivalent disposition; scores in the 40 to 49 range refer positive disposition; and scores in the 50 to 60 refer to high disposition (Facione & Facione, 1992). The overall scores of the CCTDI range between 70 and 420 and interpreted on the basis of the following standards: A total score falls in the 70 to 209 range signifies negative disposition toward critical thinking; a total score falls in the 210 to 279 range signifies ambiguity or ambivalence toward critical thinking; and a score falls in the 280 to 420 range signifies positive disposition toward critical thinking (Facione & Facione, 1992). As explained in the test manual, although the ranges defined for interpretation of scores are considered to be universal, the ranges may also be arranged or adapted on the basis of the normative standards of any group to which the CCTDI will potentially be administered. The scoring procedures and particulars of score calculations have not been revealed due to international copyrights.

Following the development of the CCTDI, other researchers, especially psychologists showed great interest in the CCTDI. Enthusiasm in understanding the interrelation of such conceptualization with already existing concepts led to the conduct of studies seeking correlations between the CCTDI and other already available research instruments and constructs such as “openness to experience” (Costa & McCrae, 1985), "ego-resiliency" (Block & Block, 1980). Sánchez (1993) found positive correlations between the scales of the CCTDI and ego resiliency: Systematicity ($r=.47$, $N=200$, $p<.001$), Truth Seeking ($r=.41$, $N=200$, $p<.001$), and Inquisitiveness ($r=.39$, $N=200$, $p<.001$); as well as with the openness to

experience construct: Truth-Seeking ($r=.27$, $p<.001$), Open-mindedness ($r=.33$, $p<.001$), CT Self-Confidence ($r=.25$, $p<.004$), Inquisitiveness ($r=.37$, $p<.001$), and Cognitive Maturity ($r=.30$, $p<.001$).

The U.S. Department of Education carried out a research on the assessment of critical thinking dispositions of students and reviewed all of the inventories available in terms of several criteria. As a result of this study U.S. Department of Education revealed a national report (2000), entitled *Definitions and Assessment Methods for Critical Thinking, Problem Solving, and Writing*, included inventories and their specifications. The report also indicated that there are only a few instruments developed to measure critical thinking and none of these instruments are designed to measure critical thinking disposition and professional judgment in teacher education except for the CCTDI, which is well conceptualized and developed to measure the extent to which a person possesses the characteristics of the ideal critical thinker.

Authorized Arabic, Chinese (Mandarin), Dutch, Farsi, Finnish, French (Canadian), Hebrew, Italian, Japanese, Korean, Portuguese, Spanish (Mexico-Latin America), and Thai language versions of the CCTDI is currently available. As mentioned in the former parts of the current section, the English version of the CCTDI displayed positive psychometric properties for use with English speaking American populations. However, the other language versions of the CCTDI were not tested in terms of their psychometric properties, except for Chinese version by Yeh (2002). Yeh reported positive alpha coefficients for two of the sub-scales of the Chinese version of the CCTDI: (Inquisitiveness, 0.73 & self-confidence, 0.68). For other sub-scales, the alpha coefficients ranged between 0.34 and 0.47, which were below 0.50 and considered to be refined with further developmental actions. In

addition, Yeh found that the measurement model hypothesized by Facione (1990) did not fit the Chinese data well. The results of his study suggested further adaptation and validation to use all of the dispositional dimensions of the CCTDI with Chinese samples. Insight Assessment which is a division of California Academic Press attained the researcher of this dissertation as the Authorized translator to produce the Turkish language version of the CCTDI and analyze the psychometrics for Turkish and American samples for Cross-cultural validation.

3.4 Translation and Back-Translation Process

Prior to any translation attempt, all the necessary permissions (see Appendix A) to translate the CCTDI have been obtained from Insight Assessment, which is a division of California Academic Press and the copyright holder of the instrument, and the author of this dissertation was attained as the authorized translator and copyright holder of the Turkish version of the CCTDI . The author of this dissertation was competent and fluent in both source and target languages and cultures with English language literature and critical pedagogy backgrounds. In addition, Hambleton (2005) and Brislin (1970) support that the researcher can be attained as the main translator in a test adaptation process if the researcher has background knowledge in the phenomenon subjected to translation. Especially Brislin suggests that in such research projects the initial translation should be carried out by the researcher if he/she is cable translating from source to target language. The reason Brislin asserts for his indication was related to the fact that the researcher him/herself is the only person who knows the construct being translated better than anyone else.

The first attempt in undertaking the process of translation was to clarify the actual intention of each item in the original CCTDI. Having an extensive informed

opinion regarding the intention of each item prior to an initial translation was quite important for maximizing semantic, conceptual, and normative equivalencies and minimizing item bias across language versions (Ægisdóttir, Gerstein, & Çinarbas, 2007; Behling & Law, 2000). For this reason, every single item in the inventory has been negotiated with Peter Facione, who is the author of the inventory, and obtained detailed informed opinion regarding the intended meaning of each item in the original CCTDI. Following meaning clarification, initial translation and back-translation process took place. For the current study, translation and back-translation process, as suggested by Brislin (1970), was embedded to an interactive adaptation process, to maximize the linguistic equivalency, in which each cycle involved three important steps: (1) Initial translation, (2) Back-translation, (3) Comparison of the original and back-translated versions for any modification and adaptation. For this reason, the following multiple interactive translation process was used to produce a linguistically equivalent Turkish version of the CCTDI (see Figure 3.4.1.).

Cycle 1 – Step 1: The author of this dissertation translated the CCTDI from English source language to Turkish target language (English Version 1 to Turkish version 1).

Cycle 1 – Step 2: A second translator, who is a bilingual and has a background in regard to the field of study was professionally hired, back translated Turkish version 1 to English version 2 without any knowledge of the English version 1.

Cycle 1 – Step 3: English version 2 back-translation was then compared to English version 1 by Peter A. Facione and a panel attained by California Academic Press (Developers and Copyright Holders of the Instrument).

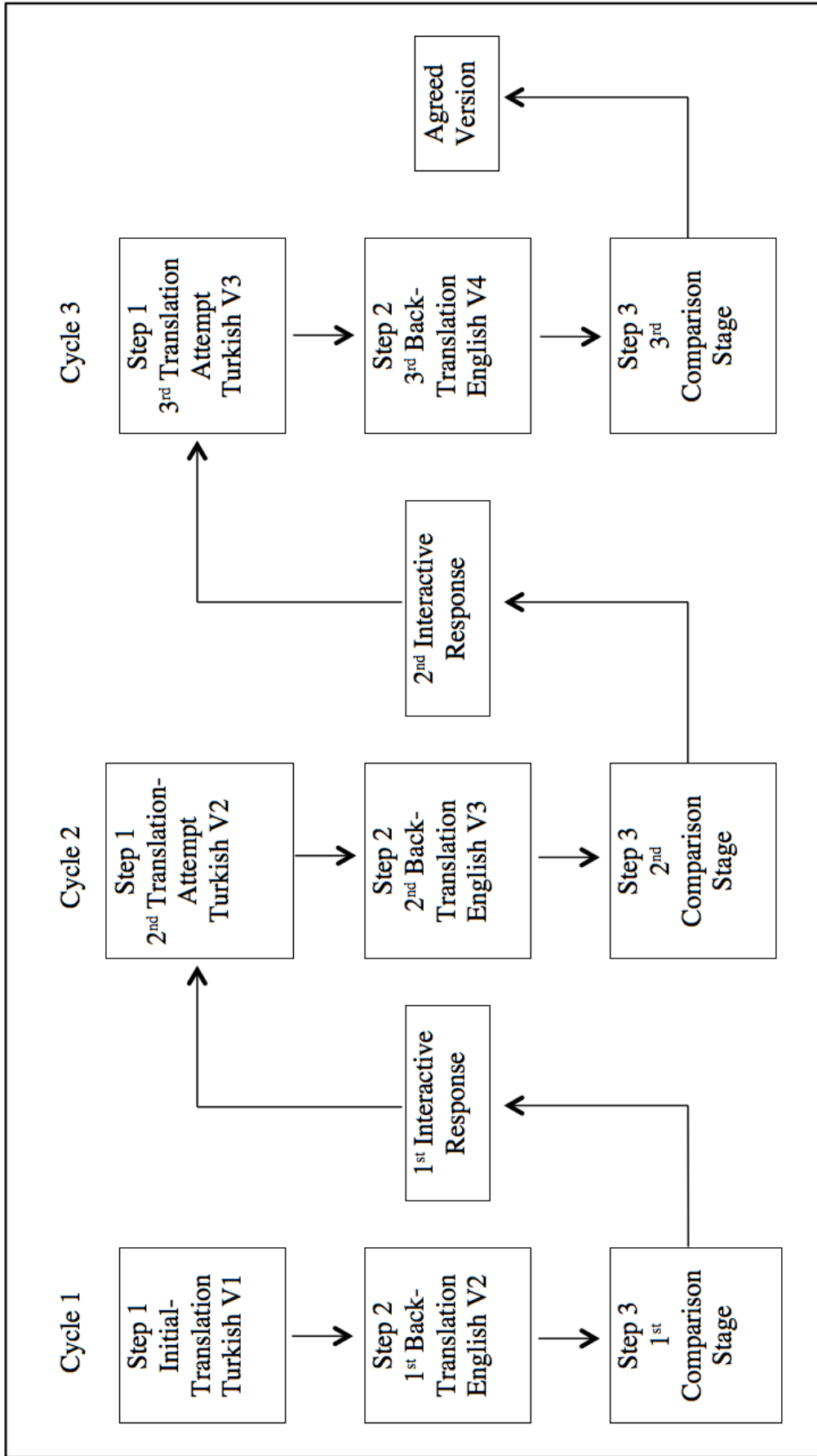


Figure 3.4.1. Multiple Interactive Translation and Back-translation Process

1st Interactive Response: As a result of the 1st comparison stage, reviewers sent a list of items that need to be changed, repaired, or responded to. According to the response, 17 items have been found problematic in terms of linguistic equivalency; thus, the intention and/or the actual message was not found in those specified 17 items in the English version 2 back-translation. On the basis of the 1st response received from California Academic Press (CAP), each of the 17 items has been negotiated online in order to detect the root of the problem. As a result of this attempt, it was concluded that some items in the Turkish version didn't include the intended message because of lexical preference, and the actual intentions of several items could not have been merged into Turkish target language because of normative, conceptual, and semantic problems. In addition, some items that included proverbs did not exist in Turkish culture. Therefore, these items need to be adapted to Turkish culture with the condition of protecting their original intentions (Weeks, Swerissen, & Belfrage, 2007).

Cycle 2 – Step 1: Regarding the decisions drawn from the first interactive response, the author has made the necessary adaptations and has re-written specified 17 items by considering the nuances in Turkish culture with special attention to preserve the original intention, and came up with Turkish version 2.

Cycle 2 – Step 2: As the related literature suggests (Herrera, DelCampo, & Ames, 1993; Weeks, Swerissen, & Belfrage, 2007) each time a different independent translator should be attained to proceed with a new translation of the instrument. For this reason, a third independent translator, who is also bilingual and studied in the field of higher order thinking, back translated the Turkish version 2 to English version 3 without any knowledge of the English version 1 and Turkish version 1.

Cycle 2 – Step 3: English version 3 back-translation sent to the panel of experts in CAP to be compared to English version 1 for any linguistic bias.

2nd Interactive Response: Second translation attempt was successful in terms of reducing the number of problematic items from 17 to 3. Those 3 items were still the English proverbs. Modifications were not enough to come up with Turkish items that conveyed the original intentions of the English items. As negotiated the situation with Peter A. Facione and the expert panel, we decided to totally change these items in such a way that would both preserve the original message and be a Turkish proverb at the same time.

Cycle 3 – Step 1: On the basis of the decision made during the second interactive response stage, attention has been directed to produce 3 new Turkish items that could be considered identical with the original English items in terms of the construct being measured and were also in the form of proverbs. Though, this attempt required more time and effort in order to meet the conditions. We knew that each of these three items intended to measure a different facet of critical thinking disposition; therefore, the construct that each item intended to measure was different in terms of the message they conveyed. For this reason, 15 new optional items in the form of proverbs that might be considered as an alternative were created. In doing so, several factors such as the original items, the meaning clarification report sent by CAP, the experience gathered from the interactive response stages, and the theory based sources regarding the phenomenon being translated have been considered at the same time (Behling & Law, 2000). Prior to ending up with a third Turkish version, the author selected the best 3 items among the many items generated for each single construct. As a result of this process, 3rd Turkish version of CCTDI has

been emerged under the light of the experiences drawn from this multiple interactive translation process.

Cycle 3 – Step 2: As the cross-control check step, the latest Turkish version 3 was back translated to English version 4 by a different independent bilingual translator, who is expert in translating proverbs, without any knowledge of the original text and previous Turkish versions.

Cycle 3 – Step 3: The English version 4 back-translation sent to CAP for further evaluation of those particular three items. This was the last interactive comparison stage run with CAP. As a result of their evaluation of those three items, they agreed with the Turkish version 3 as the identical Turkish version of the CCTDI.

For the current research; however, finalizing this translation and adaptation phase was not enough to prove that the agreed Turkish version of the CCTDI was Cross-culturally valid. Translation of inventories for Cross-cultural use need to be followed by Cross-cultural equivalency study to make sure what is being measured exists and functionally equivalent across cultures (Brislin, 1970; DiStefano & Hess, 2005; Nunnally & Bernstein; 1994).

3.5 Sampling and Data Collection Procedures

For the purpose of the current dissertation study, the original English version and the translated Turkish version of the CCTDI needed to be administered to a representative number of American and Turkish participants to ensure for psychometric properties, to check for functional equivalency, and to explore for comparability of the CCTDI scores across Turkish and American samples. Therefore, the Turkish CCTDI was decided to be administered to Turkish participants in Turkey and the original English version of the CCTDI was decided to be administered to

American participants in the United States. Thus, Hambleton (2005) supported to administer the dual language versions of measurement instruments in their natural settings with their native speakers of the language of the tests.

For the American sample, the study was presented to department chairs of 15 Universities in the United States. 3 universities gave permission to collaborate for the current research, with the condition of succeeding from an online course offered to international researchers. This online course was all about ethics and politics of carrying out research in the United States, offered by Collaborative International Training Initiative (CITI), included modules on vulnerable human resources and historical perspectives, and lasted for three months. The researcher took an online exam after completing each module. The condition was to get at least 80 out of 100 to pass each module and the online course was composed of 11 distinct modules. After completing all of the modules successfully and receiving the Human Subjects Research Curriculum Report, research was introduced to the potential participants in those 3 universities and informed them about the confidentiality of their answers and asked for permission via consent form to participate in the study. The names of the universities and participants were not mentioned anywhere in this study as they were confidential and required not to be mentioned. The same consent form was translated into Turkish version and the study introduced to potential participants in a Turkish university.

For the purposes of the current adaptation study, judgmental participant selection procedure was employed. The concern in regard to Cross-cultural judgmental sampling design is based on putting attention to only controlling key variables (Hambleton, 2005; Sekaran, 1983). The key variables represent the major

characteristics of the pilot and/or validation sample(s) (Sekaran, 1983). The key variables for both samples were their native language, culture, and their major of study. In other words, being enrolled in a teacher education program, being a native speaker of the language of the inventory, and identifying oneself as Turkish or American were enough to be a participant for this study. However, in order to determine the subjects that were representative of the central tendencies of the nation and the culture, extra effort has been put to judge the subjects in terms of several criteria such as the place of birth, the amount of time spent in the nation, languages they speak, and reasons for identifying themselves as Turkish or American.

The corresponding literature suggests to reach at least 400 ($n \geq 400$) participants for each language and culture versions to be validated (Hu & Bentler, 1999; Tucker et al., 2006). Especially Tucker et al., stated that some tests such as chi-square test, confirmatory factor analysis tests, and invariance tests are sensitive to sample size and suggested at least 400 ($n \geq 400$) participants for Cross-cultural comparison of hypothetical assessment structures and adaptation studies.

Based on the standards and guidelines suggested above, a Turkish sample of 583 ($n=583$) freshmen to senior undergraduate students (51,3% female) from 5 different teacher education programs was obtained in Turkey. Another sample of 448 ($n=448$) freshmen to senior undergraduate students (54,7% female) was obtained from 5 different teacher education programs in the United States. The Turkish sample included 231 (39.6%) freshman, 179 (30,7%) sophomore, 124 (21,3%) junior, and 49 (8,4%) senior students, whose ages ranged from 18 to 29, with a mean age of 19 ($SD = 1,72$). The American sample, which was diverse in terms of ethnicity but relatively equal in terms of program status (see Table 3.5.1.), included 121 (27,0%) freshmen,

116 (25,9%) sophomore, 106 (23,7%) junior, and 105 (23,4%) senior students, whose ages ranged from 19 to 56, with a mean age of 21 (SD = 4,04).

Table 3.5.1. Demographic Characteristics of Participants

Characteristics	American Sample		Turkish Sample	
	n	(%)	n	(%)
Ethnicity				
African American	106	(23.7)	0	(0.0)
Anglo American, Caucasian	161	(35.9)	0	(0.0)
Asian American	87	(19.4)	0	(0.0)
Hispanic, Latino, Mexican	61	(13.6)	0	(0.0)
Native American	33	(7.4)	0	(0.0)
Turkish	0	(0.0)	583	(100)
Gender				
Female	245	(54.7)	299	(51.3)
Male	203	(45.3)	284	(48.7)
Age				
18	0	(0.0)	202	(34.6)
19	147	(32.8)	178	(30.5)
20	95	(21.2)	111	(19.0)
21	89	(19.9)	43	(7.4)
22	64	(14.3)	21	(3.6)
23	11	(2.5)	18	(3.1)
24	10	(2.2)	2	(0.3)
25	3	(0.7)	0	(0.0)
26-35	21	(4.7)	8	(1.4)
36-56	8	(1.8)	0	(0.0)
Major				
Science Education	84	(18.8)	0	(0.0)
Adult Education	96	(21.4)	0	(0.0)
Special Education	87	(19.4)	0	(0.0)
Math Education	81	(18.1)	0	(0.0)
History Education	100	(22.3)	0	(0.0)
English Language Education	0	(0.0)	75	(12.9)
Computer Education	0	(0.0)	92	(15.8)
Elementary Education	0	(0.0)	114	(19.6)
Preschool Education	0	(0.0)	186	(31.9)
Sociology Education	0	(0.0)	116	(19.9)
Program Status				
Freshmen	121	(27.0)	231	(39.6)
Sophomore	116	(25.9)	179	(30.7)
Junior	106	(23.7)	124	(21.3)
Senior	105	(23.4)	49	(8.4)

Administering the CCTDI to Turkish participants required 15 separate sessions and lasted for 2 months whereas administering the CCTDI to American participants

at 3 different Universities required 46 sessions and required 8 months to complete. Nearly one year of time has been given to administer the CCTDI to both samples. The administration and completion of the CCTDI for each session took approximately 20 minutes.

The mode of administration and standards had already been identified by Facione and Facione (1992) in the test manual of the CCTDI. So, this test manual was considered to be the standardized guideline for test administration. For the purpose of administration, these guidelines were shared with the research partners in the research sites and informed them about the sensitivity of the issue that the way the CCTDI administered to different participants should depend on the same standards at every session. As Anastasi and Urbina (1997) argue, the way the same test administered to different participants at different times and different places should be administered with the same standards in order to conclude that the differences between the variables due to true differences between the participants rather than the mode of administration. Therefore, prior to administration, the professional personal hired for administering the CCTDI was provided with an online training about how to administer the CCTDI.

3.6 Procedures for Analyzing the Psychometric Properties

Following the translation and back-translation process, the first step was to check the psychometric properties of both linguistic versions of the CCTDI. As suggested by Nunnally and Bernstein (1994), when an inventory was adapted from one language to another, content validity, construct validity, internal consistency reliability and test-retest reliability should be investigated for the targeted version of the inventory. However, in order to check the psychometric equivalency between the

linguistic versions, both translated Turkish and original English language versions of the CCTDI were assessed in terms of their psychometric properties.

The first attempt in undertaking the issue was to assess the content validity of both Turkish and English versions of the CCTDI. This was required to ensure whether the items adequately measure the content domains that hypothetically desired to measure (Grant & Davis, 1997; Nunnally & Bernstein, 1994). The actual process that required content experts to judge each item against their definitions was composed of a set of events that need to followed respectively. For that mater, the first issue was to select the content experts. Experience in their fields of study, history of publications, research on the phenomenon of critical thinking, and qualifications have been considered as criteria in selecting content experts (Grant & Davis, 1997; Davis, 1992; Drasgow & Probst, 2005; Hambleton, 2005; Polit & Beck, 2006). After three months of negotiations with several Universities in Turkey and in United States, 5 Turkish experts from Turkey and 5 American experts from U.S. in educational psychology, linguistic, and critical pedagogy accepted to serve as content validators for the translated Turkish version and the original English version of the CCTDI respectively. Each expert was provided with a set of 4 different documents (Doc 1: Cover Letter, Doc 2: Content Domains, Doc 3: the CCTDI, Doc 4: Content Validity Estimation Scale). The first document, which is ‘cover letter’, informed content validators about the study, confidentiality of their answers, their roles as content validators, measurement model of the CCTDI, and detailed information of other 3 attached documents. The second document, which included the content domains and element definitions from which all 75 items of the CCTDI were obtained, served as a standard for content validators to compare each item against the

definition. In order for experts to see the complete inventory, translated Turkish version of the CCTDI (see Appendix B) was given to Turkish experts and original English version was given to American experts. Experts were also given the content validity index (CVI) developed by Waltz, Strickland, and Lenz (1991) to rate each item of the CCTDI on their relevance to content domains (ranging from 1 = not relevant to 4 = very relevant), clarity (ranging from 1 = not clear to 4 = very clear), simplicity (ranging from 1 = not simple to 4 = very simple), and ambiguity (ranging from 1 = doubtful to 4 = meaning is clear) on the four-point scale. The index for accepting a sub-scale or a total instrument valid to the specified content was the percentage of sub-scale items or overall scale items considered to be valid by receiving a score of 3 or 4 from each category (Drasgow & Probst, 2005; & Waltz et al., 1991) with a minimum content validity index of .90 (Davis 1992; Yaghmaie, 2003; Wynd, Schmidt, & Schaefer, 2003).

After examining the CVIs of both versions via expert agreements, alpha coefficients for the seven sub-scales of the CCTDI were computed with both American and Turkish samples to assess the internal consistency reliability prior to any confirmatory factor analysis (CFA) and invariance analysis (IA) attempt. Computing coefficient alphas for scales of the CCTDI before CFA and after CFA with modified items was necessary to check if any item trimming led unacceptable decrement in alpha coefficients (Nunnally & Bernstein, 1994; Yang & Green, 2011). Since this was the first attempt to translate the CCTDI from English source language to Turkish target language, the cutoff point, as recommended, for internal consistency reliability considered to be greater than .70 (George & Mallery, 2003; Gliem & Gliem, 2003; & Yang & Green, 2011).

A group of participants ($n = 53$) from the actual Turkish sample ($n = 583$) and another group of participants ($n = 38$) from the actual American sample ($n = 448$) were asked to take the test in two months apart to examine the time interval test-retest reliability. Pearson's correlation coefficient r was calculated to determine the extent to which the two sets of scores (Time 1 Scores and Time 2 Scores) were correlated for both Turkish ($n = 53$) and American ($n = 38$) data sets. In addition, paired sample t-test analysis for comparing the mean scores and F tests for testing the equivalency of variance have been used. The most important part in this phase of psychometric analysis was the interval put between the first time and the second time the test was administered to the same participants. The time interval between the occasions should not have been too long to let the construct change naturally in participants or too short to make them remember their previous responses to the items in the CCTDI (Fraenkel & Wallen, 2006). In order not to violate this rule, two months of time were allocated for the time interval. Two months of time were not enough to make a significant difference with the critical thinking dispositions of the participants but were more than enough to let them forget the responses they made for the items on the CCTDI. As Metzler and Blankenship (2008) hypothesized, for a construct such as critical thinking disposition to be changed in an individual from one direction to another (from positive to negative or vice versa), one needs to be exposed to a significant change in his/her life spaces, life conditions and standards, and perceptions for a considerably long time. Disposition is defined as a characterological profile or a habit of mind; for this very reason, scholars do not expect a dramatic change in the intellectual functioning of individuals unless they were affected by an external force which could also be considered as a series of

planned actions (Benesch, 1993; Ernst & Monroe, 2006; Facione, Facione, & Giancarlo, 2000).

3.7 Procedures for Analyzing the Cross-cultural Factorial Validity

Remembering again, the issue of adapting psychological tests into a different language and culture required a decentering approach, meaning that both versions were subject to any necessary modification or refinement (Hambleton, 2005). For that matter, the factorial validity of both English and Turkish versions of the CCTDI was examined.

Confirmatory factor analysis (CFA) is the most advanced and the strongest way to provide evidence for Cross-cultural construct validity of dual language versions of an inventory (AERA, APA, & NCME, 1999; Hambleton, 2005; Kline, 2005; Sireci, Yang, Harter, & Ehrlich, 2006). For this very reason, in order to test the degree of existence of the hypothesized 7 factor unidimensional measurement model of the CCTDI (see Figure 3.7.1) across Turkish and American samples, CFA, which is considered to be the most advanced technique for testing hypotheses about measurement models (Kline, 2005), was employed. According to Kline, the primer requirement to utilize CFA for any adaptation and validation research is to make sure if the development of inventory has a strong theoretical basis. For that matter, CFA was found to be the most appropriate analysis technique for the current dissertation study since the current study aimed to adapt and validate the CCTDI, which was originally developed in English language for use with American culture and has a strong theoretical ground hypothesizing the 7 factor measurement model (see Figure 3.7.1) with empirical evidence supporting this structure.

As it is clear with the name of the analysis, the aim is to confirm the extend of the existence of the hypothesized measurement model in the targeted samples of corresponding cultures and languages. Since there is more than one sample, the sort of analysis is also labeled as “multiple sample confirmatory factor analysis” (Kline, 2005; Nunnally & Bernstein, 1994; Schreiber, Nora, Stage, Barlow, & King, 2006). In order to better understand the basic assumptions of CFA, the following guideline

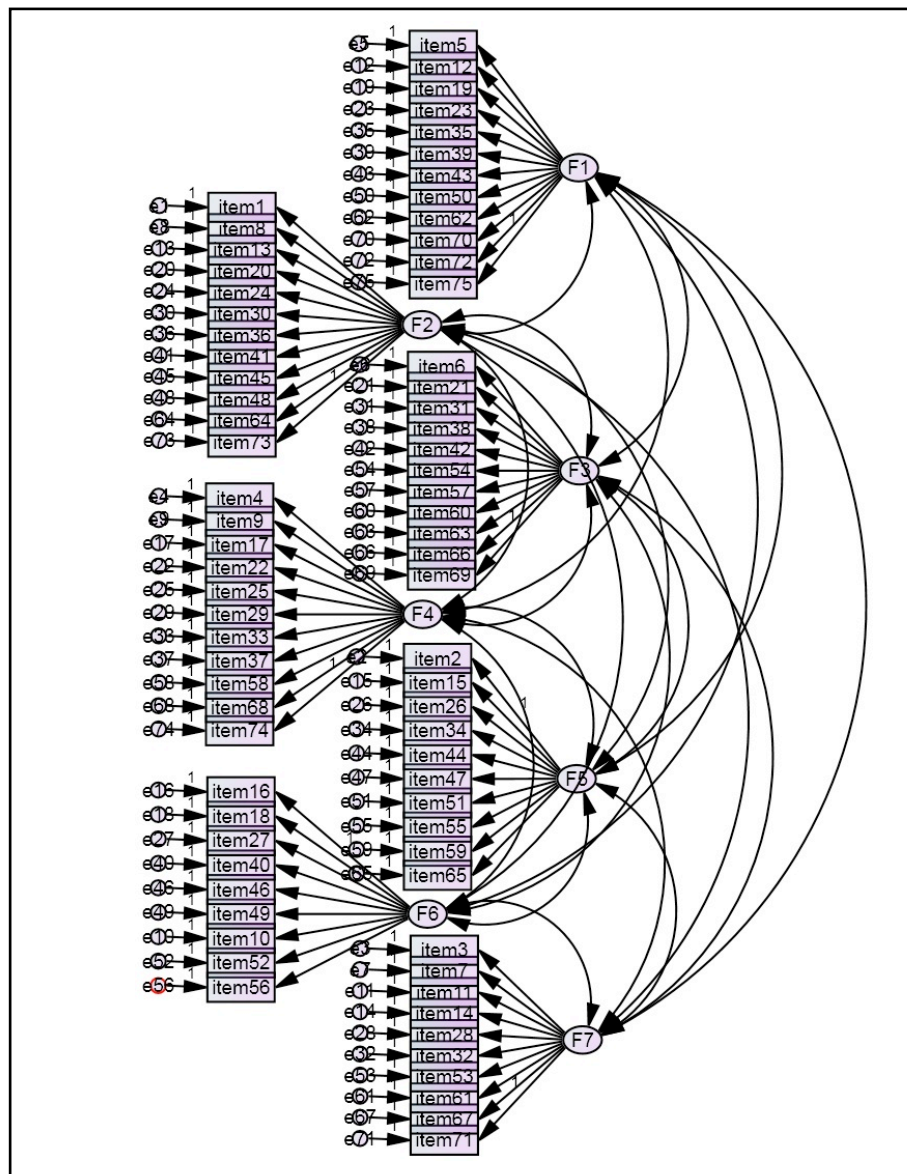


Figure 3.7.1. Hypothesized Unidimensional Seven-factor Measurement Model of the CCTDI

Note: F1: Truth-Seeking; F2: Open-Mindedness; F3: Analyticity; F4: Systematicity; F5: Inquisitiveness; F6: Critical Thinking Self Confidence; F7: Maturity of Judgement.

is suggested by Kline (2005): If the hypothesized 7 factor measurement model is correct for both samples, the analysis of results should yield less than .85 estimated correlations between the indicators of each factor. Apart from that, if the model is correct, then, each indicator, which intends to measure an underlying factor, should possess standardized loadings preferably higher than .30. The former refers to the discriminant validity if the correlations between factors are smaller than .85 and the later refers to convergent validity. To conduct CFA, Pearson product moment correlations were preferred instead of polychoric correlations because polychoric correlations were found to be non-practical in improving the model fit during the nested model analysis stage (Chen, 2007; Sun, 2005). The current research preferred to use the following indices and cutoff points to evaluate the model fit of seven factor model of CCTDI on the basis of the related literature (Chen, 2007; Hu & Bentler, 1999; Milfont & Fisher, 2010; Sass, 2011): chi-square (χ^2), degrees of freedom (df), chi-square to degrees of freedom ratio ($\chi^2/df < 4.0$), the root mean square error of approximation (RMSEA $< .06$ = good fit; values between .06 and .08 as adequate fit; and values between .08 and .10 as mediocre fit), the standardized root mean square residual (SRMR $< .06$ = good fit; values between .06 and .08 as adequate fit; and values between .08 and .10 as mediocre fit), and comparative fit index (CFI $\geq .90$ = adequate fit; and values greater than .95 as good fit).

3.8 Procedures for Analyzing the Cross-cultural Measurement

Invariance

The evaluation of the existence of the 7-factor structure of CCTDI via initial confirmatory factor analysis do not provide evidence for the extend of comparability of mean scores across cultural groups (multiple sample analysis). The actual concern

here is to test the measurement invariance, which is defined by Kline (2005) as “whether a set of indicators assess the same construct in a same way across different groups” (p. 295). In order for a researcher to use the different language versions of the same instrument to make reliable and valid Cross-cultural mean comparisons of the related construct, measurement invariance of the instrument should be tested (Behling & Law, 2000; Cheung & Rensvold, 1999; Chen, Sousa, & West, 2005; Hambleton, 2005; Kline, 2005; Sass & Schmitt, 2010; & Tucker et al., 2006). Kline clearly explains the typical practice of measurement invariance analysis and as he puts it, the analysis involves “the comparison of the relative fits with the χ^2 difference statistic of models, one with cross-group equality constraints imposed on some of its parameters and the others without constraints” (p. 295).

For testing measurement invariance across cultural groups, the criteria suggested by Chen (2007) were designated. According to Sass (2011), to develop an accurate judgment regarding the invariance model fit, Δ RMSEA, Δ SRMR, and Δ CFI should be considered in addition to ($\Delta\chi^2$) statistic; because, the use of ($\Delta\chi^2$) statistic alone may mislead and is very sensitive to large sample sizes with complex measurement models. Therefore, more practical criteria for accepting an invariance model fit involved: Δ RMSEA \leq .015, Δ SRMR \leq .03, and Δ CFI \leq .01 for interpreting the results for tests of factor loading invariance and Δ RMSEA \leq .015, Δ SRMR \leq .01, and Δ CFI \leq .01 for interpreting the results for tests of intercept invariance and residual invariance (Chen, 2007; Sass, 2011). We used SPSS 18 version together with IBM AMOS 20 version to run the required statistical analysis throughout the study.

3.9 Limitations of the Study

The current dissertation study has several limitations. First, the study is limited in its scope. That is, the purposes were only to adapt the CCTDI from English source language and culture to Turkish target language and culture; assess the psychometric properties of both of the Turkish and the English versions of the CCTDI; and explore the Cross-cultural comparability of the results obtained from both Turkish and American samples.

Second limitation is about the literature being reviewed. Researcher only focussed on the studies that were necessarily correlated with the scope and the variables being explored for the dissertation study. Although the studies carried out in the field of critical thinking and teacher education were not limited by those mentioned in the literature review chapter, the current study only placed the ones that were critically associated with the scope of the research such as studies about definitions and conceptualizations, approaches to nurture critical thinking for teacher candidates, assessment of critical thinking, and Cross-cultural research.

The third limitation is related to the limitation of population generalizability. The participants were 583 Turkish students from various teacher education programs of a Turkish university and 448 American students from various teacher education programs of an American university. Depending on the aim and the characteristics of the samples of this dissertation study, the results obtained from Turkish sample can only be generalized to Turkish speaking prospective teachers in Turkish higher education institutions and the results obtained from American sample can only be generalized to English speaking prospective teachers in American higher education institutions. However; even though the characteristics of the participants within each

language may be different from one another, consideration of the fact that the common element of representativeness of the validation samples was the language of participants yielded that the Turkish version of the CCTDI can be administered to adolescents in various disciplines in Turkish higher education institutions, and the same is valid for the English version as well.

From methodological point of view, this study is limited with empirical Cross-cultural psychometric methodology, which limited the study with the sort of statistical techniques being utilized and the results derived from the research.

Finally, the fifth type of limitation is about the measurement models hypothesized as a result of the measurement invariance tests. The resulted additional measurement models for the CCTDI are not limited by themselves and they are only the ones generated from the available validation samples. Many other models can be generated by different combinations and there are limitless number of cases in the universe but this process requires additional theoretical and empirical support for each of the alternative models that can possibly be developed. That is to say, the absence of other alternative models, other than the ones negotiated in this study, is considered to be a limitation as well.

3.10 Summary

This chapter explained the reader the procedures and the rationale behind how has been done what needs to be done to achieve the objectives of the research with supportive statements from the related literature. The researcher in the current study, who employed a cross-sectional descriptive empirical research design that was supported by the Cross-cultural measurement theory and psychometric theory to adapt and Cross-culturally validate the CCTDI, used SPSS 18 version together with

IBM AMOS 20 version to run the required statistical tests and analysis throughout the study.

Chapter 4

RESULTS

4.1 Introduction

Based on the samples collected from American and Turkish populations, two cultural data sets were used to study the psychometric properties, Cross-cultural factorial equivalency, and measurement invariance across English and Turkish versions of the CCTDI. The cut off points and other suggested criteria for assessing the Cross-cultural equivalency have been projected in great detail in a separate section in the previous chapter (Chapter 3 - Methodology) of this dissertation study. Readers should refer to this chapter for more detailed information regarding the procedures of the following statistical and qualitative analysis of the results. This chapter, in particular, was dedicated to outline the results of the analysis. Respectively, the suggested adaptations as a result of the Cross-cultural translation-back translation process, which is qualitative in its nature; the results of reliability and validity analysis regarding to both language versions of the CCTDI; Cross-cultural factorial validity across Turkish and American samples; and measurement invariance test results have been documented. In addition, pre-service teachers' dispositions of toward the seven facets of the CCTDI were analyzed and documented for both cultural groups as well.

4.2 Linguistic Equivalency across Language Versions

Adaptation of a psychological instrument for use with another culture and language involves many steps and procedures. One of the steps of such an adaptation study is the translation-back translation process that merges items into a target language and culture. Achieving a linguistically equivalent and/or identical version of any psychological test faces researchers to deal with semantic, conceptual and normative issues. In this regard, the current research study, which primarily aimed to end up with a culturally non-biased and linguistically equivalent CCTDI for use with Turkish speaking adolescents in Turkish higher education institutions, especially with prospective teachers, considered to answer the following first research question.

Research Question 1: Given findings regarding the translation and back-translation process, which adaptations are required to end up with a linguistically equivalent Turkish version of the CCTDI?

The translation-back translation technique, suggested by Brislin (1970), was embedded to an interactive adaptation process, to maximize the linguistic equivalency, in which each cycle involved three important steps: (1) Initial translation, (2) Back-translation, (3) Comparison of the original and back-translated versions for any modification and adaptation. In overall, the suggested cycle has been repeated three times to end up with a culturally non-biased and linguistically equivalent Turkish version of the CCTDI. In specific, the result of the initial translation phase revealed that 17 items for the first Turkish version of the CCTDI were found to be semantically, conceptually and normatively non-equivalent (see Table 4.2.1.), which suggested several adaptations for each of these items.

Table 4.2.1. Distribution of Problematic Items across the Types of Linguistic Equivalency

Problematic Items	Types of Linguistic Equivalency		
	Semantic	Conceptual	Normative
Item 2	✓	✗	✗
Item 3	✗	✓	✓
Item 9	✓	✗	✗
Item 18	✓	✗	✗
Item 24	✓	✗	✗
Item 37	✓	✗	✗
Item 48	✓	✗	✗
Item 50	✓	✗	✗
Item 53	✗	✓	✓
Item 57	✓	✗	✗
Item 58	✓	✗	✗
Item 68	✓	✗	✗
Item 70	✗	✗	✓
Item 71	✓	✗	✗
Item 72	✓	✗	✗
Item 73	✓	✗	✗
Item 75	✓	✗	✗

As displayed in the above table (see table 4.2.1.), item 2, item 9, item 18, item 24, item 37, item 48, item 50, item 57, item 58, item 68, item 71, item 72, item 73, and item 75 were found to be non-equivalent as a result of the linguistic check undertaken by the committee attained by the California Academic Press (CAP) on the basis of the back-translated versions. The major source of non-equivalency was due to semantic problems. The choice of words and phrases for those 17 items of the Turkish version had yielded deviations in the meanings that each item conveyed in the back-translated versions. Depending on the further negotiations of the intentions of each item, a second and third translations and back-translations were produced by different translators at different times and occasions. Here, the focus of attention was intentionally directed on the choice of words only. Lexical preferences had significant effects on changing the intention, meaning, or the message of the items.

For item 2, for example, the choice of word had significantly changed the actual message the original item had. The original statement for item 2 was: “Studying new things all my life would be wonderful”. This item was first rendered into Turkish language as: “Hayatım boyunca çalışarak yeni şeyler öğrenmek harika olurdu”. When an independent translator attempted to back-translate this Turkish statement into English without any knowledge of the original item, the statement conveyed a different meaning in English: “Working and learning new things all of life would be wonderful”. Here, when this statement was compared to the original statement by the committee in CAP, they easily notified the deviation in the actual meaning. As Facione, in his response, included, this particular item was not about “working” or “searching” new things but rather was about “studying” or “learning new things”. For this reason, a second and third translation and back-translation attempts were processed. When the choice of words and phrases were modified, the result was fascinating. The translation: “Tüm hayatım boyunca yeni şeyler öğrenmek harika olurdu” was back translated as: “Learning new things all my life would be wonderful”. Depending on this statement the Turkish statement was accepted as linguistically equivalent to the original statement in the CCTDI. This led to preserve the item for the Turkish CCTDI too. Similar processes were repeated for each of the 14 items that were found to be semantically problematic in terms of linguistic equivalency.

The second source of error that created linguistic non-equivalency between the source and target language versions were the conceptual and normative problems that were identified for Item 3 and item 53 (see Table 4.2.1.). Item 53, for instance, was not only an analogy (i.e., metaphor) but also it was talking about analogies. In

other words, this particular item housed a meaning in analogy by using an analogical example in it. The original statement, which was “Analogies are about as useful as a sailboat on a freeway”, intended to tell the test taker that analogies are not useful. But, this opinion was tested by making an analogy from western culture. Likewise, when the statement was directly translated into Turkish target language, the translation: “Analojiler (metaforlar) sadece otoban kenerına bırakılmış tekneler kadar kullanışlıdır” does not make sense for Turkish adolescents since they don't have the same nuances in Turkish culture. Here, the nuance is about the sailboats on a freeway. In U.S. people leave their old and useless sailboats on a freeway for recycling. Analogies in the original statement were valued the same as those sailboats left on a freeway. However, it is impossible for Turkish people to grasp the meaning and get the message from this illustrative analogical statement. Therefore, when this item directly translated into Turkish language, it would have created conceptual and normative problems in terms of linguistic equivalency. In order to solve this problem for this particular item, the statement was re-written with a conceptually and normatively equivalent example in it. The modified Turkish version, which was: “Benzetmeler ancak karada yüzücü paleti ile yürümek kadar kullanışlıdır”, was back-translated as “Analogies are just about as useful as walking with a swimmer-flipper on the road” and accepted by CAP as an identical Turkish version of this particular item since it conveyed the original message and intention to assess the value given to analogies. A similar process was repeated for the item 3, which was also considered to have the fingerprints of conceptual and normative problems as sources of errors for not achieving linguistic equivalency.

Item 70; however, was unique in those 17 items regarding the source of error it maintained to the investigation. It was the only item that created only normative problem. Normative non-equivalency existed when the manner in expressing the ideas are different (Hui & Triandis, 1983). In this sort of situations, Cross-cultural research is so sensitive to keep track of those normative differences or similarities to end up with a normatively equivalent version of an instrument (Asner-Self & Marotta, 2005). Item 70, which was: “I know what I think, so why should I pretend to ponder my choices”, included a norm and manner that is accepted by western culture. In the case a test taker does not agree with this statement, which is a possibility, this means that the person is pretending to ponder his/her choices as a result of his/her characterological profile. However, delivering this message to a Turkish test taker in exactly the same manner could have possibly lead him/her to avoid to tell the truth or respond the item in an expected way. For that matter, the item was negotiated with Peter Facione and the committee attained by the CAP. As a result of continuous elaborations on the issue, the item was translated as “Kafamda bir fikir oluşmuşsa, seçenekleri değerlendiriyor gibi davranmama gerek yoktur” and back-translated as “If I have a constant idea about an issue, it is unnecessary to pretend to ponder alternatives”. Since the manner of delivery of the message was changed, the translation maintained normative equivalency by preserving the original intention of the item.

In sum, as a result of the multi-cycled translation-back translation process, the above adaptations were suggested for ensuring about the linguistic equivalency of the Turkish version of the CCTDI. Although the process achieved an agreed Turkish version of the CCTDI, the adaptation process was not ended here. This was only the

first phase of the adaptation. Empirical evidence regarding the psychometric properties of both of the linguistic versions of the CCTDI was necessary to support the adaptations and modifications suggested during the translation-back translation process.

4.3 Psychometric Properties of the CCTDI across Language Versions

Following translation and back translation process, the second phase of the adaptation was to assess the psychometric properties of Turkish and English versions of the CCTDI. Since decentering approach has been utilized, both versions of the inventory were subject to necessary adaptations. In order to test the psychometrics of the CCTDI for both samples, the following research question was addressed:

Research Question 2: Given findings regarding the necessary statistical analysis, what do both the translated Turkish and the original English versions of the CCTDI demonstrate in terms of their psychometric properties?

The standard deviations and means for each item as well as for sub-scales were computed to notice the central tendencies prior to content validity analysis (see Table 4.3.1.). Following that the first psychometric check was carried out for the content validities for both Turkish and English language versions of the CCTDI. The goal was to see the extent to which each item represent the corresponding latent factor's content domains. The CVIs ranged from 0.83 to 0.99 for the sub-scales of the Turkish CCTDI and ranged from 0.97 to 1 for the sub-scales of the English CCTDI. Evidence for content validity in the sub-scales existed across versions. Although the raters agreed that the items were relevant to the content domains and definitions specified for each corresponding latent factor, lower agreement was found for the open mindedness scale in the Turkish version (see Table 4.3.2.).

Table 4.3.1. Means and Standard Deviations for Turkish and American Samples

Item No (Related Factor)	Turkish Sample		American Sample	
	Mean	SD	Mean	SD
Item 1 (OP)	4.14	1.50	3.92	1.53
Item 2 (IQ)	4.92	1.39	4.45	1.54
Item 3 (MJ)	2.87	1.40	3.07	1.52
Item 4 (SYS)	3.08	1.57	3.25	1.58
Item 5 (TS)	2.91	1.59	3.25	1.57
Item 6 (AN)	4.50	1.57	4.12	1.60
Item 7 (MJ)	2.77	1.56	3.03	1.44
Item 8 (OP)	4.10	1.41	3.92	1.40
Item 9 (SYS)	4.48	1.51	4.13	1.56
Item 10 (CT)	4.39	1.26	4.23	1.31
Item 11 (MJ)	3.03	1.43	3.22	1.47
Item 12 (TS)	3.75	1.52	3.64	1.40
Item 13 (OP)	3.53	1.83	3.65	1.70
Item 14 (MJ)	3.31	1.68	3.43	1.67
Item 15 (IQ)	3.67	1.65	3.74	1.54
Item 16 (CT)	4.73	1.57	4.27	1.60
Item 17 (SYS)	3.96	1.59	3.71	1.54
Item 18 (CT)	3.63	1.38	3.63	1.32
Item 19 (TS)	3.55	1.59	3.60	1.51
Item 20 (OP)	4.11	1.69	4.06	1.59
Item 21 (AN)	4.62	1.55	4.21	1.58
Item 22 (SYS)	4.00	1.37	3.81	1.33
Item 23 (TS)	2.98	1.48	3.82	1.40
Item 24 (OP)	3.14	1.63	3.48	1.57
Item 25 (SYS)	3.81	1.65	3.54	1.59
Item 26 (IQ)	4.77	1.23	4.41	1.31
Item 27 (CT)	4.08	1.28	3.82	1.21
Item 28 (MJ)	4.27	1.67	4.04	1.63
Item 29 (SYS)	2.87	1.54	3.13	1.51
Item 30 (OP)	4.70	1.49	4.35	1.55
Item 31 (AN)	4.70	1.25	4.25	1.36
Item 32 (MJ)	4.41	1.57	4.19	1.51
Item 33 (SYS)	3.86	1.66	3.63	1.61
Item 34 (IQ)	4.19	1.60	4.06	1.48
Item 35 (TS)	3.28	1.55	3.41	1.47
Item 36 (OP)	3.89	1.54	3.90	1.41
Item 37 (SYS)	3.93	1.51	3.81	1.48
Item 38 (AN)	4.49	1.44	4.10	1.52
Item 39 (TS)	3.27	1.45	3.33	1.33
Item 40 (CT)	4.43	1.29	4.24	1.37

(Table Continued)

Table 4.3.1. Means and Standard Deviations for Turkish and American Samples

Item No (Related Factor)	Turkish Sample		American Sample	
	Mean	SD	Mean	SD
Item 41 (OP)	4.00	1.40	3.91	1.39
Item 42 (AN)	3.61	1.49	3.60	1.45
Item 43 (TS)	2.93	1.70	3.18	1.67
Item 44 (IQ)	4.51	1.47	4.35	1.48
Item 45 (OP)	3.61	1.54	3.62	1.42
Item 46 (CT)	3.88	1.29	3.80	1.29
Item 47 (IQ)	3.98	1.42	3.83	1.32
Item 48 (OP)	3.87	1.45	3.75	1.38
Item 49 (CT)	4.32	1.35	4.08	1.44
Item 50 (TS)	3.59	1.45	3.70	1.34
Item 51 (IQ)	4.07	1.46	3.81	1.51
Item 52 (CT)	4.51	1.17	4.10	1.37
Item 53 (MJ)	3.58	1.49	3.44	1.49
Item 54 (AN)	4.34	1.35	3.82	1.45
Item 55 (IQ)	4.65	1.26	4.09	1.48
Item 56 (CT)	4.19	1.14	3.85	1.33
Item 57 (AN)	4.97	1.33	4.30	1.64
Item 58 (SYS)	3.99	1.66	3.71	1.65
Item 59 (IQ)	4.27	1.40	3.85	1.56
Item 60 (AN)	4.22	1.70	3.74	1.75
Item 61 (MJ)	4.60	1.52	3.96	1.65
Item 62 (TS)	4.25	1.46	3.95	1.46
Item 63 (AN)	3.91	1.24	3.65	1.38
Item 64 (OP)	4.41	1.51	3.87	1.62
Item 65 (IQ)	4.77	1.47	4.05	1.63
Item 66 (AN)	3.36	1.62	3.36	1.56
Item 67 (MJ)	4.28	1.76	3.73	1.81
Item 68 (SYS)	3.37	1.58	3.34	1.63
Item 69 (AN)	3.68	1.25	3.31	1.36
Item 70 (TS)	3.17	1.45	3.25	1.43
Item 71 (MJ)	3.12	1.71	3.17	1.64
Item 72 (TS)	3.60	1.37	3.67	1.33
Item 73 (OP)	4.18	1.50	3.85	1.50
Item 74 (SYS)	3.87	1.51	3.64	1.56
Item 75 (TS)	3.21	1.46	3.35	1.43
TS	3.37	0.92	3.47	0.95
OP	3.97	0.90	3.86	0.96
AN	4.22	1.02	3.85	1.09
SYS	3.75	0.99	3.61	1.06
IQ	4.38	0.96	4.07	0.99
CT	4.24	0.94	4.01	0.97
MJ	3.63	0.96	3.53	1.03

Note: TS = Truth seeking; OP = Open mindedness; AN = Analyticity; SYS = Systematicity; IQ = Inquisitiveness; CT = Critical thinking self confidence; MJ = Maturity of judgment; Turkish sample N = 583; American sample N = 448.

When alpha coefficients were studied with 75 items prior to a CFA attempt, alphas for the sub-scales ranged from .81 to .90 for the Turkish CCTDI and ranged from .85 to .91 for the English CCTDI (see Table 4.3.2.). These values satisfied the minimum expected criteria of .70 for the first attempt at translating and adapting the scales (George & Mallery, 2003). It should; however, be noted that the coefficient alphas were recalculated with the latest versions of these sub-scales after subsequent CFA for Cross-cultural comparability.

Table 4.3.2. Content Validity Indices and Alpha Coefficients for Sub-scales across Turkish and English Versions of the CCTDI

Sub-Scales (number of items)	Content Validity Indices		Alpha Coefficients	
	American (n=5)	Turkish (n=5)	American (n=448)	Turkish (n=583)
Truth-Seeking (12)	.97	.93	.88	.85
Open Mindedness (12)	.97	.83	.87	.82
Analyticity (11)	.98	.90	.91	.90
Systematicity (11)	.98	.93	.89	.86
Inquisitiveness (10)	.99	.96	.86	.86
CT Self Confidence (9)	.99	.99	.88	.88
Maturity of Judgment (10)	.98	.96	.85	.81
Overall (75)	.99	.93	.87	.87

The third type of reliability analyzed for both cultural samples was the time interval test-retest reliability. When the mean score results, Pearson correlation coefficients and the results of *t*-tests together with variance analysis were evaluated for Turkish sample (N = 53), evidence for test-retest reliability existed for the Turkish CCTDI. As can be viewed from the table (see Table 4.3.3.), all Pearson *r* were statistically significant at the 0.01 significance level, ranging from a high of 0.57 for analyticity sub-scale to a higher of 0.73 for critical thinking self-confidence sub-scale. For a more restricted analysis, the significance level was then adjusted to 0.001. In the second run; however, Pearson correlation coefficients showed no

Table 4.3.3. Correlations, Paired t-tests, and Variances for Sub-scale Scores of the Turkish CCTDI across Time

Sub-scale	Mean (SD)	r	t-test (df)	Variance	F-test (df)
<u>Truth-Seeking</u>					
Time 1	45.36(7.50)			56.31	
Time 2	43.85(7.64)	0.60 ^a	1.60(52)	58.36	2.58(51)
<u>Open-Mindedness</u>					
Time 1	42.53(6.96)			48.48	
Time 2	43.00(7.39)	0.70 ^a	- 0.62(52)	54.65	0.38(51)
<u>Analyticity</u>					
Time 1	44.89(4.15)			17.27	
Time 2	44.04(4.49)	0.57 ^a	1.54(52)	20.19	2.37(51)
<u>Systematicity</u>					
Time 1	39.77(5.44)			29.64	
Time 2	40.85(5.21)	0.59 ^a	1.62(52)	27.13	2.63(51)
<u>Inquisitiveness</u>					
Time 1	41.23(4.50)			20.22	
Time 2	40.87(4.98)	0.59 ^a	0.60(52)	24.81	0.36(51)
<u>CT-Self-Confidence</u>					
Time 1	40.28(5.99)			35.94	
Time 2	40.11(5.55)	0.73 ^a	0.29(52)	30.83	0.085(51)
<u>Maturity</u>					
Time 1	31.64(5.30)			28.08	
Time 2	32.58(6.21)	0.59 ^a	- 1.30(52)	38.56	1.71(51)
<u>Overall</u>					
Time 1	285.70(23.29)			542.37	
Time 2	285.30(25.08)	0.60 ^a	0.13(52)	629.18	0.017(51)

^a $p < 0.01$, N=53

Table 4.3.4. Correlations, Paired t-tests, and Variances for Sub-scale Scores of the English CCTDI across Time

Sub-scale	Mean (SD)	r	t-test (df)	Variance	F-test (df)
<u>Truth-Seeking</u>					
Time 1	29.81(6.49)			34.71	
Time 2	31.17(6.74)	0.68 ^a	0.69(37)	18.96	2.78(36)
<u>Open-Mindedness</u>					
Time 1	39.00(3.82)			22.32	
Time 2	38.02(4.02)	0.71 ^a	-0.65(37)	24.30	0.36(36)
<u>Analyticity</u>					
Time 1	41.54(3.43)			17.27	
Time 2	40.31(3.83)	0.52 ^a	1.66(37)	25.19	1.62(36)
<u>Systematicity</u>					
Time 1	37.53(5.56)			20.60	
Time 2	35.87(4.13)	0.42 ^a	2.67(37)	23.36	0.67(36)
<u>Inquisitiveness</u>					
Time 1	47.66(6.13)			25.22	
Time 2	40.87(4.89)	0.50 ^a	2.40(37)	33.76	1.24(36)
<u>CT-Self-Confidence</u>					
Time 1	41.39(5.66)			44.78	
Time 2	40.52(5.94)	0.79 ^a	0.99(37)	47.33	0.33(36)
<u>Maturity</u>					
Time 1	37.53(3.08)			23.38	
Time 2	34.11(4.02)	0.65 ^a	-2.66(37)	23.21	0.03(36)
<u>Overall</u>					
Time 1	274.46(18.56)			438.60	
Time 2	260.87(24.43)	0.84 ^a	3.76(37)	546.09	1.88(36)

^a $p < 0.01$, $N=38$

significant difference in the correlations across in time-1 and time-2 scores. The *t*-test result also showed no significant difference in the mean scores across the specified time interval. In addition to this analysis, the F test results revealed that there was an equality of variances since there was no significant difference detected in the variances of related samples. In other words, the differences observed in variances were non-significant at 0.001 significance level. Thus, a considerable degree of stability among the scores was evident between the related Turkish samples.

When the same procedures were repeated for the American sample (N = 38), evidence for test-retest reliability existed for the American CCTDI as well. As can be seen from the table (see Table 4.3.4.), all Pearson *r* were statistically significant at the 0.01 significance level, ranging from a high of 0.52 for analyticity sub-scale to a higher of 0.79 for critical thinking self-confidence sub-scale. Only a lower of 0.42 Pearson correlation coefficient for systematicity sub-scale was found in American sample. In overall outlook; however, significant amount of test retest reliability existed. Thus, there was no significant difference in the mean scores across time in two different occasions. Likewise, the F test results showed that there was an equality of variances since there was no significant difference detected in the variances of related American samples. That is to say, the differences observed in variances were non-significant at 0.001 significance level. Therefore, a considerable degree of variability and stability among the scores were evident between the related American samples as well.

4.4 Cross-cultural Factorial Validity of the CCTDI

In order to conduct further analysis regarding the psychometric properties of the CCTDI, the factorial validity of the CCTDI across Turkish and American samples were examined. To sketch the fit of the hypothesized measurement model of the CCTDI for both cultural groups, the following research question was addressed:

Research Question 3: Given findings regarding the confirmatory factor analysis (CFA), what is the extent to which the data derived from both the Turkish sample and the American sample explained the hypothesized 7-factor measurement model of the CCTDI.

When the standardized estimates were considered for the examination of factorial validity, the hypothesized seven-factor measurement model produced quite poor fit for Turkish sample, $\chi^2(df = 2679) = 10090.724$, $p < .0001$, $\chi^2/df = 3.767$, RMSEA = .069, SRMR = .096, CFI = .66, as well as for American sample, $\chi^2(df = 2679) = 10566.346$, $p < .0001$, $\chi^2/df = 3.944$, RMSEA = .081, SRMR = .100, CFI = .61. When the regression slopes and the correlation matrix were examined to find out the reason behind this poor model fit, the modification index suggested to exclude three factors from the measurement model, namely “Open-Mindedness”, “Analyticity”, and “Inquisitiveness” because of their quite low estimation effects of parameter estimates such as factor loadings and factor pattern coefficients. When those selected factors removed from the model, the adapted four-factor model produced better but still insufficient improvement for both the Turkish sample, $\chi^2(df = 813) = 3019.200$, $p < .0001$, $\chi^2/df = 3.714$, RMSEA = .068, SRMR = .083, CFI = .76, and the American sample, $\chi^2(df = 813) = 3279.212$, $p < .0001$, $\chi^2/df = 4.033$, RMSEA = .082, SRMR = .093, CFI = .72. The fit of four-factor model to the Turkish

sample was slightly better in comparison to American sample but still did not meet the criteria of a good model fit.

Further consideration of the standardized estimates for both samples revealed that there were several items with quite low standardized factor loadings smaller than .30 indicating that these items might not belong to the corresponding hypothesized latent factors. For this reason, with the intention of increasing the factorial validity for both versions, 5 items from truth seeking sub-scale (item12, item19, item23, item50, & item62), 5 items from systematicity sub-scale (item4, item29, item37, item58, & item68), 4 items from critical-thinking self confidence sub-scale (item10, item16, item18, item56), and 6 items from maturity of judgment sub-scale (item3, item7, item11, item14, item53, & item71) with factor loadings smaller than .30 were deleted. After this item deletion process, the modified hypothesized model displayed a significant but still an insufficient improvement for both American sample, $\chi^2(df = 203) = 682.324$, $p < .0001$, $\chi^2/df = 3.361$, RMSEA = .073, SRMR = .072, CFI = .91 and Turkish sample, $\chi^2(df = 203) = 730.348$, $p < .0001$, $\chi^2/df = 3.598$, RMSEA = .067, SRMR = .060, CFI = .92. When the modification indices were carefully evaluated, it was discovered that correlating 3 residuals with their pairs (see Figure 4.4.1.) produced an acceptable adequate fit for American sample and a good fit for Turkish sample (see Table 4.5.1.), which, in turn, produced a better baseline to establish a better Configural model for subsequent nested model comparison and further invariance analysis. The Pearson moment product correlations, means, and standard deviations (see Appendix C) for both cultural groups supported the suggested four-factor model configuration.

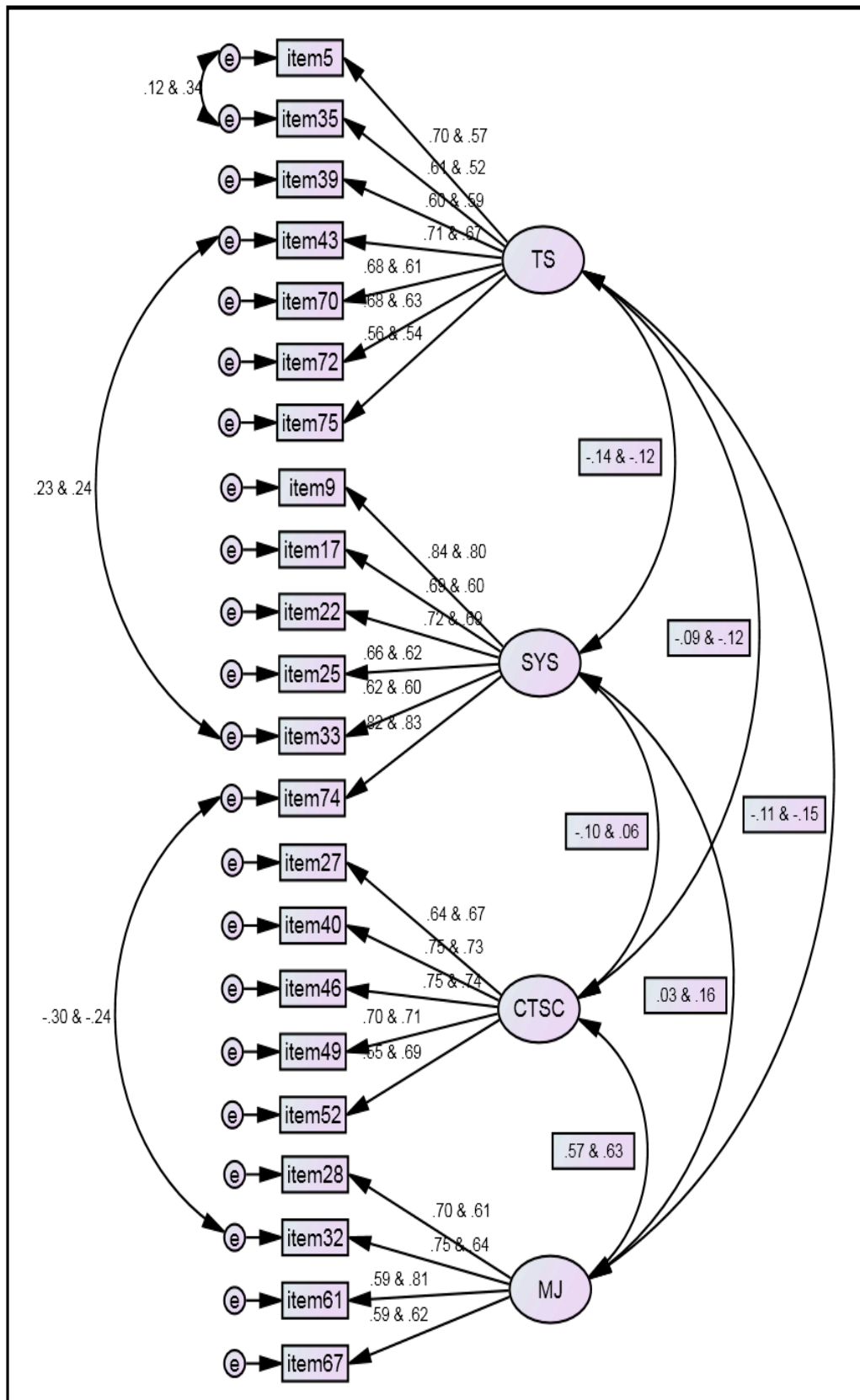


Figure 4.4.1. Modified Four-factor Model of the CCTDI with Parameter Estimates for Each Cultural Group Separately

Note: First numbers always refer to American sample. TS: Truth-Seeking Scale; SYS: Systematicity Scale; CTSC: Critical-Thinking Self Confidence Scale; MJ: Maturity of Judgment Scale.

Table 4.4.1. Specifications of the Adapted Turkish Version of the CCTDI

Sub-scale (number of items)	Items	n	FL	Mean(SD)	α	CVIs	Range
Truth-Seeking (7)	5	583	0.57	2.91(1.59)	0.80	0.93	1-6
	35		0.52	3.28(1.55)			
	39		0.59	3.27(1.45)			
	43		0.67	2.93(1.70)			
	70		0.61	3.17(1.45)			
	72		0.63	3.60(1.37)			
	75		0.54	3.21(1.46)			
Systematicity (6)	9	583	0.80	4.48(1.51)	0.75	0.92	1-6
	17		0.60	3.96(1.59)			
	22		0.69	4.00(1.37)			
	25		0.62	3.81(1.65)			
	33		0.60	3.86(1.66)			
	74		0.83	3.87(1.51)			
	27		0.67	4.08(1.28)			
CT-Self-Confidence (5)	40	583	0.73	4.43(1.29)	0.83	1	1-6
	46		0.74	3.88(1.29)			
	49		0.71	4.32(1.35)			
	52		0.69	4.51(1.17)			
	28		0.61	4.27(1.67)			
Maturity of Judgement (4)	32	583	0.64	4.41(1.57)	0.77	1	1-6
	61		0.81	4.60(1.52)			
	67		0.62	4.28(1.76)			

When the alpha coefficients were re-computed after CFA with four-factor model, the alphas for the sub-scales of Turkish and English versions of the CCTDI were respectively as follows (Truth-seeking, 7 items, $\alpha = .80$; Systematicity, 6 items, $\alpha = .75$; CT Self-confidence, 5 items, $\alpha = .83$; Maturity of judgment, 4 items, $\alpha = .77$) and (Truth-seeking, 7 items, $\alpha = .84$; Systematicity, 6 items, $\alpha = .87$; CT Self-confidence, 5 items, $\alpha = .81$; Maturity of judgment, 4 items, $\alpha = .75$). The associated table provided detailed specifications of the adapted four-factor Turkish version of the CCTDI (see table 4.4.1.). The modification suggested as a result of CFA produced decrements in alphas of all scales across cultural groups. The most decrement among the scales of English CCTDI was recorded in maturity of judgment scale and for the scales of Turkish CCTDI was recorded in systematicity scale and maturity of judgment scale. All of the alphas for the sub-scales across both cultural groups were above the critical point of .70 and displayed evidence of stability for internal consistency reliability.

4.5 Cross-cultural Measurement Invariance of the CCTDI

As the final stage of the Cross-cultural adaptation process, the degree of measurement invariance across cultural groups would be tested. Although this was one of the most complicated statistical analysis, it was highly necessary to provide evidence regarding the comparability of the results that may possibly be available by means of obtaining scores from both cultures using the CCTDI for further inferential statistical analysis. Since this was one of the purposes of the current dissertation study, the following fourth research question was answered:

Research Question 4: Given findings regarding the measurement invariance tests, what is the extent to which the translated Turkish and the original English versions of the CCTDI allow for Cross-cultural mean comparison of the construct?

The first model tested was the configural invariance model (CIM) (see Table 4.5.1.), which produced a good fit. An examination of modification indices for each cultural group separately showed that there were no residuals for any items with large modification index so that correlating residuals would not result in a significant improvement in model fit. Therefore, the initial CIM was served as a baseline model to which the imposition of more restrictive models could then be tested. The next step was to test for metric invariance model (MIM) by constraining factor pattern coefficients to be equal across cultural groups. The standard sequence for identification of non-invariant items was based on covariance matrices rather. In this regard, results regarding the comparison of Model 2 = MIM to Model 1 = CIM indicated that constraining the factor loadings across the groups achieved metric invariance from both statistical $\Delta\chi^2$ perspective and practical $\Delta\text{RMSEA} \leq .015$, $\Delta\text{SRMR} \leq .03$, and $\Delta\text{CFI} \leq .03$ perspectives (see Model 2).

Once the metric invariance model (MIM), which was a prerequisite, was supported, the scalar invariance model (SIM) was then tested. Here, we set not only the factor loadings but also the item intercepts to be equal across groups. The comparison of Model 3 = SIM to Model 2 = MIM produced a statistically significant $\Delta\chi^2$ (see Table 4.5.1.). Although Model 3 = SIM seemed to be an acceptable model from practical perspective, statistically significant $\Delta\chi^2$ meant that item parameters were unequal across groups and that led to consider partial scalar invariance model (PSIM). Vandenberg and Lance (2000) advocated that configural invariance and

Table 4.5.1. Model Fit Statistics across Cultural Groups

Model	χ^2	df	χ^2/df	$\Delta\chi^2$	Δdf	RMSEA	$\Delta RMSEA$	SRMR	$\Delta SRMR$	CFI	ΔCFI
American	636.592	200	3.183			.070		.070		.949	
Turkish	626.811	200	3.134			.061		.057		.951	
M1: CIM	1263.403	400	3.159			.046		.057		.950	
M2: MIM	1285.525	418	3.075	22.12	18	.045	.001	.058	.001	.950	.000
M3: SIM	1381.119	440	3.139	95.59*	22	.046	.001	.058	.000	.941	-.009
M4: PSIM _(i75)	1344.265	438	3.069	58.74*	20	.045	.000	.057	-.001	.951	.001
M5: PSIM _(i75 & i17)	1318.705	436	3.024	33.18*	18	.045	.000	.058	.000	.951	.001
M6: PSIM _(i75 & i17 & i25)	1311.505	434	3.021	25.98	16	.046	.001	.058	.000	.950	.000
M7: RIM	1344.745	462	3.250	33.24	22	.047	.001	.059	.001	.929	-.021
M8: FVIM	1346.325	466	3.226	1.58	4	.047	.000	.060	.001	.929	.000

Note: Statistically significant $\Delta\chi^2$ at the .050 p level were marked with an *. RMSEA = Root Mean Square Error Approximation; SRMR = Standardized Root Mean Square Residual; CFI = Comparative Fit Index; CIM = Configurational Invariance Model; MIM = Metric Invariance Model; SIM = Scalar Invariance Model; PSIM = Partial Scalar Invariance Model; RIM = Residual Invariance Model; FVIM = Factor Variance Invariance Model.

metric invariance should be satisfied in order to proceed with any further partial invariance models. Once this requirement was met, to identify the items, which might cause the model misfit, modification indices were evaluated carefully with the intention of coming up with a non-significant $\Delta\chi^2$ statistic when compared to Model 2 = MIM to produce an acceptable PSIM. In order to produce a non-significant $\Delta\chi^2$ for the Model 4 = PSIM_(i75), consideration of modification indices suggested to relax the constraints put on factor loadings and item intercepts only for item 75. Although results regarding $\Delta RMSEA \leq .015$, $\Delta SRMR \leq .01$, and $\Delta CFI \leq .01$ displayed evidence for invariance across groups, relaxing item 75 did not reveal a non-significant $\Delta\chi^2$ for Model 4, indicating that the null hypothesis of no significant differences across cultural groups could then be rejected. According to modification index, the other two items seemed to be non-invariant were the item 17 and item 25. We first preferred to relax item 17 together with item 75 and realized a significant but still an insufficient decrement in the $\Delta\chi^2$ for the Model 5 = PSIM_(i75 & i17). The factor loadings and item intercepts for the item 75, item 17, and item 25 were then relaxed and Model 6 = PSIM_(i75 & i17 & i25) revealed a good model fit with a non-significant $\Delta\chi^2$ statistic comparing to Model 2 = MIM (see Table 4.5.1.), indicating that partial scalar invariance achieved across cultural groups when those specified three non-invariant items were relaxed. Even though achieving metric and scalar invariance was considered to be enough for supporting Cross-cultural comparability of scores for inferential statistics (Milfont & Fisher, 2010; Tucker, Ozer, Lyubomirsky, & Boehm, 2006), we also tested for more restrictive invariance models. As can be seen from table (see Table 4.5.1.), we constrained error variances to produce residual invariance model (RIM) and factor variances to produce factor

variance invariance model (FVIM) respectively, in addition to factor loadings and item intercepts. Model 7 = RIM revealed statistically non-significant $\Delta\chi^2$ in comparison to less restrictive partial invariance model, Model 6 = PSIM_(i75 & i17 & i25) with accepted $\Delta RMSEA \leq .015$, $\Delta SRMR \leq .01$, and $\Delta CFI \leq .01$. As expected, Model 8 = FVIM also resulted with evidence for factorial invariance across cultural groups when compared to Model 7 = RIM, thus indicating that the range of scores on the latent factors do not vary across cultural groups.

4.6 Critical Thinking Dispositions of Pre-service Teachers

As the final concern of the dissertation study, critical thinking dispositions of pre-service teachers, based on the validation samples, were examined for both Turkish and American higher education institutions. To satisfy this requirement, the following research question was considered:

Research Question 5: Given findings regarding the adapted CCTDI, what are the critical thinking dispositions of Pre-service teachers across Turkish and American higher education institutions?

Before getting into the results of analysis for the identification of critical thinking dispositions of pre-service teachers across samples, it is necessary to note that the analysis was done at the intra level rather than inter level. In other words, within analysis was preferred rather than between analysis. Mean scores from cultural groups were not tested for their statistical differences since this is beyond the scope of this dissertation. In addition, analysis was done in terms of seven dispositional dimensions of the CCTDI across samples.

The results of analysis for the seven sub-scales and the overall scale for American sample, as outlined in the table (see Table 4.6.1.), showed that pre-service

teachers were ambivalently disposed to truthseeking, systematicity, maturity of judgment, and open-mindedness and were positively disposed to analyticity, inquisitiveness, and critical thinking self-confidence. When their overall disposition toward critical thinking was computed, it was found that they were ambivalently disposed to think critically.

Table 4.6.1. Means and Standard Deviations for the Sub-scales and the Overall Scale across Turkish and American Samples

Sub-Scales (number of items)	American Sample		Turkish Sample	
	Mean	SD	Mean	SD
Truth-Seeking (12)	33.74	9.234	33.04	9.62
Open Mindedness (12)	39.73	8.999	37.74	10.19
Analyticity (11)	42.17	10.219	39.29	10.98
Systematicity (11)	37.47	9.995	35.17	11.22
Inquisitiveness (10)	43.80	9.566	41.18	10.02
CT Self Confidence (9)	42.39	9.361	39.90	9.94
Maturity of Judgment (10)	36.26	9.592	34.34	10.57
Overall (75)	275.57	33.428	260.70	36.97

Note: American Sample N = 448. Turkish Sample N = 583

Consideration of the performances of the Turkish Pre-service teachers (see Table 4.6.1.) showed that they were ambivalently disposed to truthseeking, open-mindedness, analyticity, systematicity, and maturity of judgment and were positively disposed to inquisitiveness and critical thinking self-confidence. Their overall disposition toward critical thinking was found to be ambivalent as well. In comparison, American students performed slightly better than Turkish students on all of the sub-scales and the overall scale. A major difference was only recorded for analyticity sub-scale in favor of the American students.

Chapter 5

DISCUSSION

This research, at the beginning, developed a compelling argument for the importance of identifying a culturally relevant well conceptualized measurement tool for assessing pre-service teachers' critical thinking dispositions to augment more traditional examinations of pre-service teachers' cognitive skills in critical thinking. Given the fact that a sound English language instrument from the United States, the CCTDI, was identified and was purported to be available in several non-English versions, it made sense that the CCTDI be adapted and subjected to the translation and Cross-cultural validation process outlined in this dissertation.

The first remark to make about the findings of this study is that the translation-back translation process yielded a linguistically equivalent Turkish version of the CCTDI. Yet, some intriguing points remained. For instance, it is detected that when the items possessed semantic problems, they did not possess conceptual or normative problems. If they displayed conceptual or normative problems, then the source of problem was not semantic. Why those items that displayed semantic problems did not display normative or conceptual problems? Or, why those items that displayed both normative and conceptual problems did not reveal semantic problems? Is there a categorical or hierarchic order coded for the various sources of error in the Cross-cultural translation process. The existing literature, however, does not have any explanation regarding this finding. Therefore,

the significance of such a difference should be sought since the difference could inform international test translators about how the issue of maintaining equivalency works for translation-back translation process.

The findings of this study supported that obtaining an identical target language version of the instrument in terms of linguistic equivalency with high values of Cronbach's alpha and CVIs for the sub-scales of the CCTDI did not mean that the target language version possessed good factorial validity and measurement invariance across cultures. Thus, evidence retrieved from the results of initial run with 75 items seven-factor model for factorial validity produced a poor model fit for both cultural groups with items from each sub-scale with low parameter estimates. This finding highlighted the necessity to conduct two separate analyses for the cross-validation and measurement invariance across Turkish and American populations. Interestingly, those items showed themselves during the interactive response stage of translation back-translation process and were considered to be emic-unique to the source culture. For that matter, based on the modification indices, those items were pondered to be non-invariant and removed. Removing those items from the model displayed significant but insufficient improvement in model fit for cultural groups. Looking back to the modification indices suggested correlating three residuals with their pairs, and this time the modified 22 items four-factor model revealed acceptable model fit for American sample and good fit for the Turkish sample. Here, it is necessary to indicate that the modified four-factor structure enhanced the model fit and the modification did not distort the structure for the remaining four-factor since the alpha coefficients and the number of items allocated for each latent factor remained sufficient according to the norms specified by the literature (Cheung &

Rensvold, 2000; DiStefano & Hess, 2005). In other words, evidently the remaining items still measure what they would supposedly measure.

The other intriguing finding was that the model fit statistics showed better fit for Turkish target sample rather than the American source sample. Usually, the source language version was expected to show better fit to the observed data (Dimitrov, 2010). There are several reasons explaining this situation but one of the evidences supported that the English version of the instrument was developed by Facione (1990), and no further update was considered for the English source version of the CCTDI since then. As Chomsky (2011) argues, in *Hopes and Prospects*, language, like societies, has been evolved by the circumstances of the era. As the social realities change as a result of globalization and other associated factors, culture also do change with influence on language and perception (Chomsky, 2011). In other words, globalization influences language through the screen of society and culture. Therefore, the way people perceive the phenomenon of critical thinking is different today in comparison to their perceptions of two decades ago. The disposition aspect of critical thinking was influenced by the evolution in language, culture, and perceptions of people. With respect to that finding, this research underlines the importance of considering linguistic theories when assessing equivalency between the linguistic versions of the measurement instruments, and suggests further studies to modify the items to update the original version of the CCTDI by considering the perceptions of people of the source culture and re-assess the factorial validity with a similar study. Therefore, one of the reasons for explaining better fit with Turkish sample can be considered to be the three-cycle multiple interactive translation procedures, which updated the items of the Turkish version of the instrument in

terms of language use and proverb preference. The narrative and ultimate resolution of the proverb dissonance between language translations is a vivid example of the importance of the cultural validation endeavor and a crucial contribution to the growing literature on the international applicability of assessment instruments.

Following factorial validity check, the current research also examined the measurement invariance of the CCTDI across cultural groups to see whether researchers could proceed with a meaningful Cross-cultural mean comparison. Each model was tested against more restrictive models and the results of analysis derived from the comparison achieved a full metric invariance from both statistical $\Delta\chi^2$ and practical $\Delta\text{RMSEA} = .001$, $\Delta\text{SRMR} = .001$, and $\Delta\text{CFI} = .000$ perspectives. However, the results did not support a full scalar invariance but rather showed that relaxing constraints put on item 75, item 17, and item 25 exhibited partial scalar invariance. Regardless of the messages each of these three items intend to give to the test taker, the way the items were written might have caused potential non-invariance across groups since they were constructed in the form of proverbs. It can be said that the use of proverbs in the construction of such psychological tests might be the cause of invariance lacking across cultural groups.

The issue regarding the use of proverbs in psychological tests has two dimensions. The first dimension includes the argument of whether or not to use proverbs in psychological tests and the second one includes arguments about whether or not researchers should consider translating and using these proverbs for Cross-cultural comparison of constructs (Behling & Law; 2000). A group of researchers have accumulated to indicate that the use of proverbs may be allowed to some extent if the construct being measured is emic-unique to the source culture (Behling & Law;

2000; Berry, 1969; Yang, 1997). However, this is not valid for etic constructs since etic constructs possess same components by means of definition regardless of the culture. When the content domains and definitions for latent factors of four-factor model of the CCTDI were evaluated, it is obvious to state that the constructs being measured were etic. This may therefore be considered as one of the evidences explaining the potential cause of non-invariance across cultural groups. In other words, an etic construct that intends to measure the phenomenon under investigation with emic items could not be merged successfully into target language and culture for Cross-cultural mean comparison. However, the current research does not provide strong evidence regarding why the items constructed in the form of proverbs were found to be non-invariant. Therefore, this can also be considered as a concern of a future research, which should be conducted to investigate the possible reasons behind such non-invariance. Future research is also needed to deepen understanding of other possible causes of differences, thus the differences may be due to the administration of tests, translation errors, participants' perceptions of the items, culture specific emic constructs, and different conceptions of critical thinking disposition. The analysis of results of the current research further continued that, when additional constraints employed to test more restrictive models, evidence for residual invariance and factorial invariance existed across cultural groups. Overall, support for partial scalar invariance indicated that latent means could be meaningfully compared across cultural groups without any measurement bias. However, the differences might be due to a reason which we do not know yet. A differentiated item functioning (DIF) study can also be suggested for further researchers to get deeper insight into why several items functioned for Turkish culture in a different way (Ercikan, 2002). Cultural

conventions of the Turkish society, the construct of critical thinking, and the items' relatedness to the culture need to be thought together when studying DIF for those items.

Chapter 6

CONCLUSION

The current research, which attempted to adapt the CCTDI and investigate the Cross-cultural validity for the purpose of assessing the critical thinking dispositions of Pre-service teachers across Turkish and American higher education institutions, led important contributions to the existing related literature. Consideration of the results has documented the following suggestions for researchers who would like to use the CCTDI to measure one's dispositions toward critical thinking and to utilize Cross-cultural mean comparison (Cheung & Rensvold, 1999; Milfont & Fisher, 2010). First, despite the fact that the initial 75 items seven-factor model of the CCTDI did not achieve good fit to the observed data, researchers can use this model to collect data to assess the critical thinking dispositions of Pre-service teachers in Turkish higher education institutions by relying on the high values of Cronbach's alpha for internal consistency reliability, high values of test-retest reliability, and high values of CVIs for evidence for content validity (Cheung & Rensvold, 1999). Second, researchers may omit the three specified latent factors that were considered to be non-invariant and delete the items with low parameter estimates and use the modified 22 items four-factor model to assess the disposition dimension of critical thinking across Turkish and American higher education institutions. Third, researchers may either prefer to use the partial scalar invariance model or assume that the differences between cultural groups are not big enough to

influence the results, so use all the items to proceed with a Cross-cultural mean comparison. Fourth, researchers may simply use the scales for within culture analysis or use the scales but avoid Cross-cultural mean comparisons.

In conclusion, assessing critical thinking dispositions has become a wider issue and no longer a local issue. With an increasing interest in international research, researchers seek to find reliable and valid instruments to make Cross-cultural comparisons. Using existing instruments to measure a construct in another cultural group requires Cross-cultural validation study rather than a simple translation procedure. In addition to that, this research will lead to question the Arabic, Chinese (Mandarin), Dutch, Farsi, Finnish, French (Canadian), Hebrew, Italian, Japanese, Korean, Portuguese, Spanish (Mexico-Latin America), and Thai language versions of the CCTDI since there is no study examining the factorial validity or empirical evidence for Cross-cultural applicability of those language versions except for Chinese (Mandarin) and Turkish language versions. This study also underlines an important fact that even though all the procedures are employed to translate and back-translate psychological instruments for Cross-cultural use, this does not ensure that the translated version displays factorial equivalency between the linguistic versions of the instruments. Therefore, researchers should hold with strong evidences supporting that the results regarding Cross-cultural comparisons due to real differences in the people and the variables being measured rather than errors in translation or any other reasons. One of the contemporary ways of providing evidence regarding the issue is to go with confirmatory factor analysis as well as measurement invariance analysis across cultural groups.

REFERENCES

- Adamopolous, J., & Lonner, W. J. (2001). Culture and psychology at a crossroad: Historical perspective and theoretical analysis. In D. Matsumoto (Ed.), *The handbook of culture and psychology* (pp. 11-34). New York: Oxford University Press.
- Ægisdóttir, S., Gerstein, L. H., & Çinarbas, D. C. (2007). Methodological issues in Cross-cultural counseling research: Equivalence, bias, and translations. *The Counseling Psychologist, 36*(2), 188-219. Retrieved October 21, 2008, from <http://wjn.sagepub.com/cgi/content/refs/36/2/188>.
- Allalouf, A., Hambleton, R. K., & Sireci, S. G. (1999). Identifying the causes of DIF in translated verbal items. *Journal of Educational Measurement, 36*(3), 185-198. Retrieved October 21, 2008, from the JSTOR database.
- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: Author.
- American Psychological Association. (2003). Guidelines on multicultural education, training, research, practice, and organizational change for psychologists. *American Psychologist, 58*, 377-402.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). London: Prentice Hall.
- Asner-Self, K. K., & Marotta, S. A. (2005). Developmental indices among Central

- American immigrants exposed to way-related trauma: Clinical implications for counselors. *Journal of Counseling and Development*, 83, 162-171.
- Bailin, S., Case, R. Coombs, R. J., & Daniels, L. B. (1999a). Common misconceptions of critical thinking. *Journal of Curriculum Studies*, 31(3), 269-283. Retrieved June 1, 2007, from the Taylor and Francis database.
- Bailin, S., Case, R. Coombs, R. J., & Daniels, L. B. (1999b). Conceptualizing critical thinking. *Journal of Curriculum Studies*, 31, 285-302. Retrieved June 1, 2007, from the Taylor and Francis database.
- Barton, E. P. (2009). *National education standards: Getting beneath the surface*. Retrieved from Reports in Educational Testing Service website: www.ets.org/research/pic
- Behling, O., & Law, K. S. (2000). *Translating questionnaires and other research instruments: Problems and solutions*. California: Sage Publications.
- Benesch, S. (1993). Critical thinking: A learning process for democracy. *TESOL Quarterly*, 27(3), 545-548. Retrieved June 1, 2007, from the JSTOR database.
- Benesch, S. (1999). Thinking critically, thinking dialogically. *TESOL Quarterly*, 33(3), 573-580. Retrieved June 1, 2007, from the JSTOR database.
- Bensley, D. A. (2006). Why great thinkers sometimes fail to think critically. *Skeptical Inquirer*, 30(4), 47-52.
- Berry, J. W. (1969). On Cross-cultural comparability. *International Journal of Psychology*, 4, 119-128.
- Berry, J. W. (1989). Imposed etics-emics-derived etics: The operationalization of a compelling idea. *International Journal of Psychology*, 26, 721-735.
- Betz, N. E. (2005). Enhancing research productivity in counseling psychology:

- Reactions to three perspectives. *The Counseling Psychologist*, 33, 358-366.
- Block, J. H. & Block, J. (1980). The role of ego-control and ego-resiliency in the organization of behavior. In W. A. Collins (Ed.), *Development of Cognition, Affect, and Social Relation: The Minnesota Symposia on Child Psychology*, Vol. 13. Hillsdale, NJ: Lawrence Earlbaum.
- Brislin, R. W. (1970). Back-translation for Cross-cultural research. *Journal of Cross-cultural Psychology*, 1(3), 185-216
- Brislin, R. W. (1976). Comparative research methodology: Cross cultural studies. *International Journal of Psychology*, 11, 213-229.
- Brislin, R. W. (1983). Cross cultural research in psychology. *Annual Review of Psychology*, 34, 363-400.
- Brislin, R. W. (1986). The wording and translation of research instruments. In W. J. Lonner & J. W. Berry (Eds.), *Field methods in Cross-cultural research* (pp. 137-164). Beverly Hills, CA: Sage.
- Brislin, R. W., Lonner, W. J., & Thorndike, R. M. (1973). *Cross-cultural research methods*. New York: John Wiley.
- Brookfield, S. D. (1997). Assessing critical thinking. *New Directions for Adult and Continuing Education*, 75, 17-29.
- Browne, M. N., & Freeman, K. (2000). Distinguishing features of critical thinking classrooms. *Teaching in Higher Education*, 5(3), 301-309. Retrieved June 1, 2007, from the Taylor and Francis database.
- Browne, M. N., & Keeley, S. M. (1988). Do college students know how to “think critically” when they graduate? *Research Serving Teaching*, 1(9), 2-3.
- Bucy, M. C. (2006). Encouraging critical thinking through expert panel discussions.

College Teaching, 54(2), 222-224.

Byrne, M. S., & Jhonstone, A. H. (1987). Critical thinking and science education.

Studies in Higher Education, 12(3), 325-339. Retrieved June 1, 2007, from the Taylor and Francis database.

Chapman, D. W., & Carter, J. F. (1979). Translation procedures for the Cross-cultural

use of measurement instruments. *Educational Evaluation and Policy Analysis*, 1(3), 71-76. Retrieved October 21, 2008, from the JSTOR database.

Chen, F. F. (2007). Sensitivity of goodness of fit indices to lack of measurement

invariance. *Structural Equation Modeling*, 14, 464-504.

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Testing measurement invariance of

second-order factor models. *Structural Equation Modeling*, 12, 471-492.

Cheong, A. C. S., & Loong, D. H. W. (2007). Developing a thinking culture within

the Singaporean university context. *Asia Pacific Journal of Education*, 19(2), 72-85. Retrieved June 1, 2007, from the Taylor and Francis database.

Cheung, G. W., & Rensvold, R. B. (1999). Testing factorial invariance across groups:

A reconceptualization and proposed new method. *Journal of Management*, 25(1), 1-27

Cheung, G. W., & Rensvold, R. B. (2000). Assessing extreme and acquiescence

response sets in Cross-cultural research using structural equation modeling. *Journal of Cross-cultural Psychology*, 31, 188-213.

Chomsky, N. (2011). *Hopes and prospects*. Chicago: Haymarket Books.

Coleman, H., Rogers, G., & King, J. (2002). Using portfolios to stimulate critical

thinking in social work education. *Social Work Education*, 21(5), 583-595.

Retrieved June 1, 2007, from the JSTOR database.

Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory Manual*.

Odessa, FL: Psychological Assessment Resources.

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts.

Applied Nursing Research, 5, 194-197.

Dewey, J. (1910). *How we think*. Boston: D. C. Heath.

Dewey, J. (1938). *Experience and education*. New York: Kappa Delta Pi.

Dewey, J. (2004). *Democracy and education*. United Kingdom: Kessinger Publishing's Rare Reprints (Original work published 1916).

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, 43(2) 121–149

DiStefano, C., & Hess, B. (2005). Using confirmatory factor analysis for construct validation: An empirical review. *Journal of Psychoeducational Assessment*, 23, 225-241. Retrieved October 21, 2008, from <http://wjn.sagepub.com/cgi/content/refs/23/3/225>

Douglas, N. L. (2000). Enemies of critical thinking: Lessons from social psychology research. *Reading Psychology*, 21(2), 129-144. Retrieved June 1, 2007, from the Taylor and Francis database.

Dragow, F., & Probst, T. M. (2005). The psychometrics of adaptation: Evaluating measurement equivalence across languages and cultures. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for Cross-cultural assessment* (pp. 3-38). New Jersey: Lawrence Erlbaum Associates.

- Dumka, L. E., Stoerzinger, H. D, Jackson, K. M., & Roosa, M. W. (1996). Examination of the Cross-cultural and cross-language equivalence of the parenting self-agency measure. *Family Relations*, 45(2), 216-222. Retrieved October 21, 2008, from the JSTOR database.
- Dworkin, M. S. (1959). *Dewey on education*. New York: Teachers College Press.
- Eklof, T. F. (2005). Higher mind: The method of critical thinking. *Philosophical Practice*, 1(3), 129-133. Retrieved June 1, 2007, from the Taylor and Francis database.
- Ellis, A. K. (2005). *Research on educational innovations* (4th ed.). New York: Eye On Education.
- Ennis, R. H. (1967). A concept of critical thinking: A proposed basis for research in the teaching and evaluation of critical thinking ability. In B. P. Komisar & C. J. B. Macmillan (Eds.), *Psychological Concepts in Education* (pp. 114-148). Chicago: Rand McNally.
- Ennis, R. H. (1980). A conception of rational thinking. In J. R. Coombs (Ed), *Proceedings of the Philosophy of Education Society* (pp. 3-27). Normal, IL: The Philosophy of Education Society.
- Ennis, R. H. (1985). Critical thinking and the curriculum. *Phi Kappa Phi Journal*, 65(1), 28-31.
- Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & J. R. Sternberg (Eds.), *Teaching Thinking Skills: Theory and Practice* (pp. 9-26). New York: Freeman.
- Ennis, R. H. (1991). Critical thinking: Streamlined conception. *Teaching Philosophy*, 14(1), 5-24.

- Ennis, R. H. (1993). Critical thinking assessment. *Theory into Practice*, 32(3), 179-186. Retrieved June 1, 2007, from the Taylor and Francis database.
- Ennis, R. H., & Weir, E. (1985). *The Ennis-Weir critical thinking essay test*. Pacific Grove, CA: Midwest.
- Ercikan, K. (2002). Disentangling sources of differential item functioning in multi language assessments. *International Journal of Testing*, 2(3), 199-215.
- Ernst, J. A., & Monroe, M. (2006). The effects of environment-based education on students' critical thinking skills and disposition toward critical thinking. *Environmental Education Research*, 12, 429-443. Retrieved June 1, 2007, from the Taylor and Francis database.
- European University Association (EUA). (2007a). *Creativity in higher education: Report on the EUA creativity project* (ISBN: 9789081069892). Belgium: EUA.
- European University Association (EUA). (2007b). *Lisbon declaration. Europe's universities beyond 2010* (ISBN: 9789078997030). Belgium: EUA.
- Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. Millbrae, CA: California Academic Press.
- Facione, P. A. (2006). *Critical thinking: What it is and why it counts*. Millbrae, CA: California Academic Press.
- Facione, P. A., & Facione, N. C. (1992). *California critical thinking disposition inventory test manual*. Millbrae, CA: California Academic Press.
- Facione, P. A., & Facione, N. C. (1992). *The California critical thinking disposition inventory*. Millbrae, CA: California Academic Press.

- Facione, P. A., Facione, N. C., & Giancarlo, C. A. (1997). The motivation to think in working and learning. In E. Jones (Ed.), *Preparing competent college graduates: Setting new and higher expectations for student learning* (pp. 67-79). San Francisco, CA: Jossey-Bass Publishers.
- Facione, P. A., Facione, N. C., & Giancarlo, C. A. (2000). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. *Informal Logic*, 20(1), 61-84. Retrieved June 10, 2007, from www.calpress.com.
- Facione, P. A., Giancarlo, C. A., & Facione, N. C. (1995). The disposition toward critical thinking. *Journal of General Education*, 44(1), 1-25.
- Fisher, A. (2001). *Critical thinking*. Cambridge: Cambridge University Press.
- Fogarty, R., & McTighe, J. (1993). Educating teachers for higher order thinking. The three-story intellect. *Theory into Practice*, 32(3), 161-169. Retrieved June 1, 2007, from the JSTOR database.
- Foucault, M. (1973). *The order of things: An archaeology of human sciences*. New York: Pantheon Books.
- Fraenkel, J. R., & Wallen, N. E. (2006). *How to design and evaluate research* (6th ed.). New York: McGraw-Hill.
- Freire, P. (1974). *Education for critical consciousness*. New York: Continuum.
- Gabennesch, H. (2006). Critical thinking: What is it good for? (In fact, what is it?). *Skeptical Inquirer*, 30(2), 36-41.
- Gadzella, B. M., Hogan, L., Masten, W., Stacks, J., Stephens, R., & Zascavage, V. (2001). Reliability and validity of the Watson-Glasere critical thinking appraisal-forms for different academic groups. *Journal of Instructional*

- Psychology*, 33(2), 141-143.
- Geertsen, H. R. (2003). Rethinking thinking about higher-level thinking. *Teaching Sociology*, 31(1), 1-19. Retrieved June 12, 2007, from the JSTOR database.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Boston: Allyn & Bacon.
- Giancarlo, C. A., & Facione, P. A. (2001). A look across four years at the disposition toward critical thinking among undergraduate students. *Journal of General Education*, 50(1), 29-55.
- Giroux, H. A. (2000) *Stealing Innocence. Corporate culture's war on children*, New York: Palgrave.
- Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting Cronbach's Alpha reliability coefficients for likert-type scales. Paper presented at Midwest Research to Practice Conference in Adult, Continuing, and Community Education, The Ohio State University, Columbus, OH, October 8-10, 2003.
- Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. *Research in Nursing & Health*, 20, 269-274. Retrieved October 28, 2008, the JSTOR database.
- Grauerholz, L., & Holtrop, S. B. (2003). Exploring critical sociological thinking. *Teaching Sociology*, 31(4), 485-496. Retrieved June 1, 2007, from the JSTOR database.
- Greenfield, P. M. (1997). You can't take it with you: Why ability assessments don't cross cultures. *American Psychologist*, 52, 1115-1124.
- Grosser, M. M., & Lombard, B. J. J. (2008). The relationship between culture and the

- development of critical thinking abilities of prospective teachers. *Teaching and Teacher Education*, 20, 1364-1375. Retrieved August 1, 2008, from the Elsevier database.
- Gruber, S., & Boreen, J. (2003). Teaching critical thinking: Using experience to promote teaching in middle school and college students. *Teachers and Teaching*, 9(1), 5-19. Retrieved June 1, 2007, from the Taylor and Francis database.
- Hager, P., & Kaye, M. (1991). Critical thinking ability and teacher effectiveness. *Higher Education Research & Development*, 10(2), 177-186. Retrieved February 27, 2009, from the Elsevier database.
- Hambleton, R. K. (2001). The next generation of the ITC test translation and adaptation guidelines. *European Journal of Psychological Assessment*, 17, 164-172.
- Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for Cross-cultural assessment* (pp. 3-38). New Jersey: Lawrence Erlbaum Associates.
- Hambleton, R. K., & de Jong, J. H. A. L. (2003). Advances in translating and adapting educational and psychological tests. *Language Testing*, 20, 127-134.
- Hare, W. (1979). *Open-mindedness and education*. Montreal: McGill-Queen's University Press.
- Harrison, N. (2004). Rethinking critical thinking: Indigenous students studying at university. *Teaching Education*, 15(4), 375-384. Retrieved June 1, 2007, from

the Taylor and Francis database.

Herrera, R. S., DelCampo, R. L., & Ames, M. H. (1993). A serial approach for translating family science instrumentation. *Family Relations*, 42(3), 357-360.

Retrieved October 21, 2008, from the JSTOR database.

Hsueh, K. H., Philips, L. R., Cheng, W. Y., & Picot, S. J. F. (2005). Assessing Cross-cultural equivalence through confirmatory factor analysis. *Western Journal of Nursing Research*, 27(6), 755-771. Retrieved October 21, 2008, from <http://wjn.sagepub.com/cgi/content/refs/27/6/755>

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-15.

Hui, C. H., & Triandis, H. C. (1983). Multistrategy approach to Cross-cultural research – The case of locus of control. *Journal of Cross-cultural Psychology*, 14, 65-83.

Ikuenobe, P. (2001). Teaching and assessing critical thinking abilities as outcomes in an informal logic course. *Teaching in Higher Education*, 61(1), 19-32. Retrieved July 14, 2007, from the Taylor and Francis database.

Johnson, L. E., & Reiman, A. J. (2007). Beginning teacher disposition: Examining the moral/ethical domain. *Teaching and Teacher Education*, 23, 676-687. Retrieved August 1, 2008, from the Elsevier database.

Johnson, T. P. (1998). Approaches to equivalence in Cross-cultural and cross-national survey research. In *Cross-cultural Survey Equivalence* (pp. 1-40). Retrieved from http://www.gesis.org/Publikationen/Zeitschriften/ZUMA_Nachrichten_spezial/zn-sp-3-inhalt.htm

- Jones, A. (2007). Multiplicities or manna from heaven? Critical thinking and the disciplinary context. *Australian Journal of Education*, 51(1), 84-103.
- Karadağ, E. (2011). Instruments used in doctoral dissertations in educational sciences in Turkey: Quality of research and analytical errors. *Educational Sciences: Theory and Practice*, 11(1), 330-334.
- Keeley, S. M., & Browne, M. N. (1966). How college seniors operationalize critical thinking behavior. *College Student Journal*, 20, 389-95.
- Keeley, S. M., & Browne, M. N. (1986). How college seniors operationalize critical thinking behavior. *College Student Journal*, 20, 389-95.
- Kim, U. (2001). Culture, science, and indigenous psychologies. In D. Matsumoto (Ed.), *The handbook of culture and psychology* (pp. 51-76). New York: Oxford University Press.
- Kline, R. B. (2005). Principles and practices of structural equation modeling. New York: The Guilford Press.
- Kurfiss, J. G. (1988). *Critical thinking: Theory, research, practice, and possibilities* (Tech. Rep. No. 2). Washington, D.C: Association for the Study of Higher Education.
- Lau, K. K. (1992). On the objectives of teacher-education. *CHUK Education Journal*, 20(1), 43-48. Retrieved August 1, 2008, from the Elsevier database.
- Lee, H. J. (2005). Understanding and assessing Pre-service teachers' reflective thinking. *Teaching and Teacher Education*, 21, 699-715. Retrieved August 1, 2008, from the Elsevier database.
- Lewin, K. (1935). *A Dynamic Theory of Personality: Selected Papers*. New York: McGraw-Hill.

- Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory into Practice*, 32(3), 131-137. Retrieved June 1, 2007, from the JSTOR database.
- Liston, D., & Zeichner, K. (1991). *Teacher education and the social conditions of schooling*. New York: Routledge.
- Mallik, A. (2004). *Technology and security in the 21st century: A demand-side perspective*. U.S. NY: Oxford University Press.
- McBride, R. E., Xiang, P., & Wittenburg, D. (2002). Dispositions toward critical thinking: The Pre-service teachers' perspective. *Teachers and Teaching: Theory and Practice*, 8(1), 29-40. Retrieved June 1, 2007, from the Taylor & Francis database.
- McBride, R. E., Xiang, P., Wittenburg, D., & Shen, J. (2002). An analysis of pre-service teachers' dispositions toward critical thinking: A Cross-cultural perspective. *Asia Pacific Journal of Teacher Education* , 30(2), 131-140. Retrieved June 1, 2007, from the Taylor & Francis database.
- Mcintire, S. A., & Miller, L. A. (2000). *Foundations of psychological testing*. New York: McGraw-Hill.
- McMillan, J. H. (1987). Enhancing college students' critical thinking: A review of studies. *Research in Higher Education*, 26 (1), 3-29.
- McPeck, J. E. (1981). *Critical thinking and education*. Oxford: Martin Robertson.
- Melnick, S. L., & Zeichner, K. M. (1998). Teacher education's responsibility to address diversity issues: Enhancing institutional capacity. *Theory into Practice*, 37(2), 88-95. Retrieved August 1, 2008, from the Elsevier database.
- Metzler, M. W., & Blankenship, B. T. (2008). Taking the next step: Connecting teacher education, research on teaching, and programme assessment.

Teaching and Teacher Education, 24, 1098-1111. Retrieved August 1, 2008, from the Elsevier database.

Milfont, T. L., & Fisher, R. (2010). Testing measurement invariance across groups:

Applications in Cross-cultural research. *International Journal of Psychological Research*, 3(1), 111-121.

Miri, B., Chaim, B., & Uri, Z. (2007). Purposely teaching for the promotion of

higher-order thinking skills: A case of critical thinking. *Research Science Education*, 37(3), 353-369.

Murphy, L. L., Conoley, J. C., & Impara, J. C. (Eds.). (1994). *Tests in print iv*.

Lincoln, NE: University of Nebraska Press.

Norris, S. P. (1989). Can we test validity for critical thinking? *Educational Research*,

18(9), 21-26. Retrieved June 1, 2007, from the JSTOR database.

Norris, S. P., & Ennis, R. H. (1989). *Evaluating critical thinking*. Pacific Grove, CA:

Midwest Publications.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York:

McGraw-Hill.

Paul, R. W., & Nosich, G. M. (1992). *A model for the national assessment of higher*

order thinking. Santa Rosa, CA: Foundation for Critical Thinking.

Pedersen, P. B. (2003). Culturally biased assumptions in counseling psychology. *The*

Counseling Psychologist, 31, 396-403.

Perkins, D. N. (1989). Reasoning as it is and could be: An empirical perspective. In

D. M. Topping, D. S. Cromwell, & V. N. Kobayaski (Eds.), *Thinking across cultures: Third international conference on thinking* (pp. 175-194). Hillsdale,

NJ: Erlbaum.

- Phillips, V., & Bond, C. (2004). Undergraduate's experience of critical thinking. *Higher Education Research & Development, 23*(3), 277-294. Retrieved June 1, 2007, from the Taylor and Francis database.
- Pitters, R. T., & Soden, R. (2000). Critical thinking in education: A review. *Educational Research, 42*(3), 237-249. Retrieved June 1, 2007, from the Taylor and Francis database.
- Plath, D., English, B., Connors, L., & Beveridge, A. (1999). Evaluating the outcomes of intensive critical thinking instruction for social work students. *Social Work Education: The International Journal, 18*(2), 207-217. Retrieved June 1, 2007, from the Taylor and Francis database.
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing and Health, 29*, 489-497. Retrieved October 28, 2008, the JSTOR database.
- Ponterotto, J. G., Casas, J. M., Suzuki, L. A., & Alexander, C. M. (1995). *Handbook of multicultural counseling*. Thousand Oaks, CA: Sage.
- Popper, K. (1935). *The logic of scientific discovery* (J. Freed & L. Freed, Trans.). London: Routledge.
- Prior, J. (2000). Social psychology of a learning environment and the acquisition of critical thinking skills. *Social Work Education, 19*(5), 501-511. Retrieved June 1, 2007, from the Taylor and Francis database.
- Reed, J. H. (1998). *Effect of a model for critical thinking on student achievement in primary source document analysis and interpretation, argumentative reasoning, critical thinking dispositions, and history content in a community college history course*. (Doctoral dissertation). Retrieved from <http://>

- Roberts, J. (1998). *Language teacher education*. London: Arnold.
- Rodrik, D. (1997). *Has globalization gone too far*. Washington D.C: Worldwide News Agency.
- Ryle, G. (1963). *The concept of mind*. Harmondsworth, UK: Penguin.
- Said, S. M., Adikan, F. R., Mekhilef, S., & Rahim, N. A. (2005). Implementation of the problem-based learning approach in the department of electrical engineering, University of Malaya. *European Journal of Engineering Education*, 30(1), 129-136. Retrieved June 1, 2007, from the Taylor and Francis database.
- Saito, S., Nomura, H., Noguchi, Y., & Tezuka, I. (1996). Translability of family concepts into the Japanese culture: Using the Family Environment Scale. *Family Process*, 35, 239-257
- Sánchez, C. A. (1993). An exploration of cognitive strategies and dispositions in relation to ego resiliency. Unpublished manuscript, University of California, Riverside.
- Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment*, 29(4), 347-363.
- Sass, D. A., & Schmitt, T. A. (2010). A Comparative investigation of rotation criteria within exploratory factor analysis. *Multivariate Behavioral Research*, 45, 1-33.
- Scheffler, I. (1966). Philosophical models of teaching. In I. Sheffler (Ed.), *Philosophy and Education: Modern Readings* (pp. 99-114). Boston: Allyn &

Bacon.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research, 99*(6), 323-337. Retrieved October 21, 2008, from the Elsevier database.

Sclove, R. (2010). *Reinventing technology assessment: A 21st century model*. Washington, D.C. Woodrow Wilson.

Scriven, M. (1985). Critical for survival. *National Forum, 55*, 9-12.

Sekaran, U. (1983). Methodological and theoretical issues and advancements in Cross-cultural research. *Journal of International Business Studies, 14*(2), 61-73. Retrieved October 21, 2008, from the JSTOR database.

Sellnow, D. D., & Ahlfeldt, S. L. (2005). Fostering critical thinking and teamwork skills via a problem-based learning (PBL) approach to public speaking fundamentals. *Communication Teacher, 19*(1), 33-38. Retrieved June 1, 2007, from the Taylor and Francis database.

Sen, A. (2007). *Development thinking at the beginning of the 21st century*. Paper presented at the conference of Inter-American Bank, Washington, D.C.

Shaughnessy, J. J., Zechmeister, E. B., & Zechmeister, J. S. (2003). *Research methods in psychology* (6th ed.). New York: McGraw-Hill.

Shepelak, N. J., Jackson, A. C., & Moore, V. L. (1992). Critical thinking in introductory sociology classes: A program of implementation and evaluation. *Teaching Sociology, 20*(1), 18-27. Retrieved August 4, 2007, from the JSTOR database.

Silverman, D. (2001). *Interpreting qualitative data: Methods for analyzing talk, text,*

and interaction (2nd ed.). Thousand Oaks, CA: Sage Publications.

Sireci, S. G., Yang, Y., Harter, J., & Ehrlich, E. J. (2006). Evaluating guidelines for test adaptations: A methodological analysis of translation quality. *Journal of Cross-cultural Psychology, 37*(5), 557-567. Retrieved October 21, 2008, from <http://jcc.sagepub.com/cgi/content/refs/37/5/557>.

Smith, M. K. (2006). *Globalization and the incorporation of education*. Retrieved March 1, 2006, from http://www.infed.org/biblio/globalization_and_education.htm

Stansfield, C. W. (2003). Test translation and adaptation in public education in the USA. *Language Testing, 20*(2), 189-207. Retrieved October 21, 2008, from <http://wjn.sagepub.com/cgi/content/refs/20/2/189>.

Suliman, W. A. (2006). Critical thinking and learning styles of students in conventional and accelerated programmes. *International Nursing Review, 53*, 73-79.

Sun, J. (2005). Assessing goodness of fit in confirmatory factor analysis. *Measurement and Evaluation in Counseling and Development, 37*(4), 240-257. Retrieved May 21, 2009, from Expanded Academic ASAP.

Tabachnick, B. R., & Zeichner, K. (1991). *Issues and practices in inquiry-oriented teacher education*. London: Falmer Press.

Tapper, J. (2004). Student perceptions of how critical thinking is embedded in a degree program. *Higher Education Research and Development, 23*(2), 199-222. Retrieved June 1, 2007, from the Taylor and Francis database.

Thomas, J. (1993). *Doing critical ethnography* (Vol. 26). London: Sage Publications.

Triandis, H. C. (1976). Approaches toward minimizing translation. In R. Brislin

- (Ed.), *Translation: Applications and research* (pp. 229-243). New York: Wiley/Halstead.
- Triandis, H. C. (1994). *Culture and social behavior*. New York: McGraw-Hill.
- Tucker, K. L., Ozer, D. J., Lyubomirsky, S., & Boehm, J. K. (2006). Testing for measurement invariance in the satisfaction with life scale: A comparison of Russians and North Americans. *Social Indicators Research*, 78, 341-360.
- Turkish Council of National Education. Lifelong Learning Program. (2009). *Turkish educational system and effective school indexes* (Research Report No. 142320-LLP-1-2008-1-TR-COMENIUS-CMP). Turkey: Ankara.
- Turkish Educational Association. (2010). *The transition system to secondary school and higher education institutions*. Ankara: Dumat Ofset Matbaacılık.
- Turkish Kızılay General Education Report. Department of Educational Management. (2011). *Developing modules for revising education*. Retrieved from http://www.kizilay.org.tr/egitimhizmetleri/2010_turk_kizilayi_egitim_raporu_internet.pdf
- Turkish Republic of Northern Cyprus Ministry of Education Report. (2005). *On project based learning*. TRNC: Nicosia.
- U.S. Department of Education, National Center for Education Statistics. (2000). *The NPEC sourcebook on assessment, volume 1: Definitions and assessment methods for critical thinking, problem solving, and writing* (NCES Publication No. 2000-172). Washington, DC: U.S. Government Printing Office.
- U.S. Department of Labor, Centre for Program Planning and Results. (2010). *Annual performance report* (Research Report No. 202-693-7120). U.S. Washington,

D.C. Government Printing Office.

- van de Vijver, F. J. R. (1998). Towards a theory of bias and equivalence. In ZUMA Cross-cultural Survey Equivalence (pp. 41-65). Retrieved from http://www.gesis.org/Publikationen/Zeitschriften/ZUMA_Nachrichten_spezial/zn-sp-3-inhalt.htm
- van de Vijver, F. J. R. (2001). The evolution of Cross-cultural research methods. In D. Matsumoto (Ed.), *Handbook of culture and psychology* (pp. 77-97). New York: Oxford University Press.
- van de Vijver, F. J. R. & Leung, K. (1997). *Methods and data analysis for Cross-cultural research*. Thousand Oaks, CA: Sage.
- van de Vijver, F. J. R., & Poortinga, Y. H. (1997). Towards an integrated analysis of bias in Cross-cultural assessment. *European Journal of Psychological Assessment, 13*, 29-37.
- Vandenberg, R. J., & Lance, C. E. (2000). A review of synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4-70.
- Wagner, T. (2008). *The global achievement gap: Why even our best schools don't teach the new survival skills our children need – and what can we do about it*. New York: Basic Books.
- Walker, P., & Finney, N. (1999). Skill development and critical thinking in higher education. *Teaching in Higher Education, 4*(4), 531-547. Retrieved June 1, 2007, from the Taylor and Francis database.
- Waltz, C. F., Strickland, O., & Lenz, E. (1991). *Measurement in nursing research* (2nd ed.). Philadelphia: F.A. Davis.

- Wang, W. L., Lee, H. L., & Fetzer, S. J. (2006). Challenges and strategies of instrument translation. *Western Journal of Nursing Research*, 28(3), 310-321. Retrieved October 21, 2008, from <http://wjn.sagepub.com/cgi/content/refs/28/3/310>.
- Weeks, A., Swerissen, H., & Belfrage, J. (2007). Issues, challenges, and solutions in translating study instruments. *Evaluation Review*, 31(2), 153-165. Retrieved October 21, 2008, from <http://erx.sagepub.com/cgi/content/refs/31/2/153>.
- Wessel, J., & Williams, R. (2004). Critical thinking and learning styles of students in a problem-based, master's entry-level physical therapy program. *Physiotherapy Theory and Practice*, 20(2), 79-89. Retrieved June 1, 2007, from the Taylor and Francis database.
- Wolcott, H. F. (2001). *Writing up qualitative research* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Wynd, C. A., Schmidt, B., & Schaefer, M. A. (2003). Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research*, 25(5), 508-518. Retrieved October 28, 2008, the JSTOR database.
- Yaghmaie, F. (2003). Content validity and its estimation. *Journal of Medical Education*, 3(1), 25-27. Retrieved October 28, 2008, the JSTOR database.
- Yang, K. S. (1997). Indigenizing westernized Chinese psychology. In M. H. Bond (Ed.), *Working at the interface of cultures: Eighteen lives in social science* (pp. 62-76). London: Routledge.
- Yang, Y., & Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment*, 29, 377-392.
- Yeh, M. L. (2002). Assessing the reliability and validity of the Chinese version of the

California Critical Thinking Disposition Inventory. *International Journal of Nursing Studies*, 39, 123-132. Retrieved August 1, 2008, from the Elsevier database.

Yoon, K., Schmidt, F., & Ilies, R. (2002). Cross-cultural construct validity of the five-factor model of personality among Korean employees. *Journal of Cross-cultural Psychology*, 33(3), 217-235. Retrieved October 21, 2008, from <http://wjn.sagepub.com/cgi/content/refs/33/3/217>.

Yu, D. S. F., Lee, D. T. F., & Woo, J. (2004). Issues and challenges of instrument translation. *Western Journal of Nursing Research*, 26(3), 307-320. Retrieved October 21, 2008, from <http://wjn.sagepub.com/cgi/content/refs/26/3/307>.

Zeichner, K. (2009). *Teacher education and the struggle for social justice*. New York: Routledge.

APPENDICES

Appendix A: CCTDI Authorization Translator Permission Letter



Insight Assessment

A division of California Academic Press
Measuring Critical Thinking Worldwide

October 28, 2008
Gökhan İskifoğlu (PhD Candidate)
Eastern Mediterranean University
Faculty of Education

Dear Gökhan İskifoğlu,

I am happy to confirm that you have permission from the authors and copyright holders, Dr. Noreen Facione and Dr. Peter Facione, and from the publisher, The California Academic Press / Insight Assessment, to translate the CCTDI. You are officially assigned as authorized translator for the Turkish version of the CCTDI by California Academic Press / Insight Assessment. I regret that it took a few extra days before I could send this message. It took me a bit of time to check with all the relevant parties.

On the finished translation cover toward the bottom of the page below the Insight Assessment logo, and on each subsequent page as footer, please indicate the following copyright.

“© 1992, Peter A. Facione and Noreen C. Facione. Published by The California Academic Press LLC, DBA Insight Assessment. All rights reserved worldwide. Turkish translation 2008 by Gökhan İskifoğlu. For information about the CCTDI contact Insight Assessment: www.insightassessment.com”

As is our custom when acknowledging our international translation collaborators, you may put your name on the cover below the authors' names and indicate that you are the translator.

Yours,

James Morante, Ph.D
Managing Director
The California Academic Press / Insight Assessment

217LACRUZAVE.
MILLBRAE, CA 94030

INSIGHT ASSESSMENT
www.insightassessment.com PHONE (650) 697-5628
FAX (650) 692-0141

Appendix B: Translated Turkish Version of the CCTDI

CCTDI

Tutum Ölçeđi

Dr. Peter A. Facione
Santa Clara Üniversitesi

Dr. Noreen C. Facione
California Üniversitesi, San Francisco

Çeviren: Gökhan İskifođlu

Yönergenin Başlamasını Bekleyin

© 1992, Peter A. Facione and Noreen C. Facione. Published by The California Academic Press LLC, DBA Insight Assessment. All rights reserved worldwide. Turkish translation 2008 by Gökhan İskifođlu. For information about the CCTDI contact Insight Assessment: www.insightassessment.com

CCTDI

İsim: _____ Öğrenci No: _____ Grup: _____

Soyisim: _____ Bölümünüz: _____

Cinsiyet

- Erkek
 Kız

Lisans Öğrencileri

1. Sınıf
 2. Sınıf
 3. Sınıf
 4. Sınıf

Yüksek Lisans

- Master
 Doktora

Uyruğunuz

- TC
 KKTC

Coğrafi Bölge

- Marmara Böl.
 İç Anad. Böl.
 Ege Böl.
 Akdeniz Böl.
 Karadeniz Böl.
 Doğu Anad. Böl.
 Güneydoğu Anad. Böl.

KATILMIYORUM (-)

KATILYORUM (+)

1	2	3	4	5	6
Hiç Katılmıyorum	Katılmıyorum	Kısmen Katılmıyorum	Kısmen Katılıyorum	Katılıyorum	Tamamen Katılıyorum

1. Tüm seçenekleri gözönüne almak, benim için karşılanamayacak bir lükstür.	1	2	3	4	5	6
2. Tüm hayatım boyunca yeni şeyler öğrenmek harika olurdu.	1	2	3	4	5	6
3. Bir öneri için en iyi argüman, o öneri hakkında o an hissettiklerindir.	1	2	3	4	5	6
4. Sorunum dikkatimin kolay dağılmasıdır.	1	2	3	4	5	6
5. Birbirleriyle çelişen fikirler arasında tercih yapmak çok zordur.	1	2	3	4	5	6
6. İnsanların iyi bir fikri savunmak için zayıf nedenlere güvenmeleri beni rahatsız eder.	1	2	3	4	5	6
7. Gerçek daima bakış açınıza bağlıdır.	1	2	3	4	5	6
8. Farkında olmadan önyargılara sahip olabileceğim fikri beni endişelendirir.	1	2	3	4	5	6
9. Cevap vermeye kalkışmadan önce, her zaman soruya odaklanırım.	1	2	3	4	5	6
10. Büyük bir isabetle düşünebildiğim için kendimle gurur duyuyorum.	1	2	3	4	5	6
11. Çoğu şey hakkında gerçekleri hiçbir zaman öğrenemeyiz.	1	2	3	4	5	6

CCTDI © 1992 Peter A. Faclone, Noreen C. Faclone, Hermosa Beach CA. Published by The California Academic Press LLC. Turkish translation by Gökhan İskifoğlu. All rights reserved. For information about the CCTDI and its authorized translations visit the Insight Assessment website: www.insightassessment.com

KATILMIYORUM (-)			KATILIYORUM (+)			
1	2	3	4	5	6	
Hiç Katılmıyorum	Katılmıyorum	Kısmen Katılmıyorum	Kısmen Katılıyorum	Katılıyorum	Tamamen Katılıyorum	
12. Dört lehte, bir aleyhte görüş varsa, lehte olan dört görüşe katılırım.	1	2	3	4	5	6
13. Erkekler ve kadınlar eşit derecede mantıklıdır.	1	2	3	4	5	6
14. Tavsiyenin değeri, karşılığında ödediğin bedel kadardır.	1	2	3	4	5	6
15. Üniversitedeki derslerin çoğu ilginç değildir ve yararsızdır.	1	2	3	4	5	6
16. Ezberi değil, yorum yapmayı gerektiren sınavlar benim için daha iyidir.	1	2	3	4	5	6
17. Problemlerim hakkında hiçbir çözüm üretmeden saatlerce konuşabilirim.	1	2	3	4	5	6
18. İnsanlar benim entelektüel merakımı ve sorgulayıcılığımı takdir eder.	1	2	3	4	5	6
19. Kanıtlar yanlışlığı gösterse bile, inandıklarına sıkı sıkıya bağlı kalırım.	1	2	3	4	5	6
20. Bariz bir şekilde hatalıysanız, düşüncelerinizi ortaya koymaya hakkınız yoktur.	1	2	3	4	5	6
21. Mantıklıymış gibi davranıyorum, ama değilim.	1	2	3	4	5	6
22. Düşüncelerimi düzenlemek benim için kolaydır.	1	2	3	4	5	6
23. Ben dahil herkes daima kendi ilgi alanı doğrultusunda tartışır.	1	2	3	4	5	6
24. Doğruluk ve yanlışlık sözkonusu olduğunda, açık fikirli olmanın sınırlıkları vardır.	1	2	3	4	5	6
25. Yaptığım harcamalarımın dikkatlice kaydını tutmak benim için önemlidir.	1	2	3	4	5	6
26. Önemli bir karararla yüz yüze geldiğimde, karar vermeden önce, toplayabileceğim tüm bilgileri toplarım.	1	2	3	4	5	6
27. Tarafsız yaklaşabildiğim için arkadaşlarım karar almada benden medet umarlar.	1	2	3	4	5	6
28. Açık fikirli olmak neyin doğru, neyin yanlış, olduğunu bilmemek demektir.	1	2	3	4	5	6
29. Bankalar, kredi kartı hesap ekstrelerini daha anlaşılır hale getirmelidirler.	1	2	3	4	5	6
30. Diğer insanların çeşitli konular hakkında neler düşündüklerini anlamak benim için önemlidir.	1	2	3	4	5	6
31. İnanıklarımın tümü için dayanaklarım olmalı.	1	2	3	4	5	6

KATILMIYORUM (-)			KATILIYORUM (+)			
1	2	3	4	5	6	
Hiç Katılmıyorum	Katılmıyorum	Kısmen Katılmıyorum	Kısmen Katılıyorum	Katılıyorum	Tamamen Katılıyorum	
32. Okumak, mümkün olduğunca kaçtığım bir şeydir.	1	2	3	4	5	6
33. İnsanlar çok acele karar verdiğimi söylerler.	1	2	3	4	5	6
34. Üniversitedeki zorunlu dersler vakit kaybına yol açar.	1	2	3	4	5	6
35. Gerçekten çok karmaşık bir şeyle uğraşmak zorunda kaldığımda paniklerim.	1	2	3	4	5	6
36. Yabancılar her zaman kendi kültürlerini anlamaya çalışmak yerine bizim kültürümüzü anlamaya çalışmalıdırlar.	1	2	3	4	5	6
37. İnsanlar benim karar vermeyi ağırdan aldığımı düşünürler.	1	2	3	4	5	6
38. Başkalarının görüşlerine/fikirlerine karşı çıkabilmeleri için insanların gerekçelere ihtiyaçları vardır.	1	2	3	4	5	6
39. Kendi inanışlarımı tartışırken tarafsız olmam imkansızdır.	1	2	3	4	5	6
40. Ortaya yaratıcı alternatifler koyabildiğim için kendimle gurur duyuyorum.	1	2	3	4	5	6
41. Doğruyu söylemem gerekirse, daha az yargılayıcı olmaya çalışıyorum.	1	2	3	4	5	6
42. Kendimi sık sık insanların argümanlarını/iddialarını değerlendirirken bulurum.	1	2	3	4	5	6
43. Neye inanmak istiyorsam ona inanırım.	1	2	3	4	5	6
44. Zor problemleri çözmek için uğraşmayı sürdürmek kısacası önemli değildir.	1	2	3	4	5	6
45. Fikirlerimi savunmak zorunda bırakılmamalıyım.	1	2	3	4	5	6
46. Başkaları, kararların uygulanmasında mantıklı standartların belirlenmesi için bana danışırlar.	1	2	3	4	5	6
47. Zorlayıcı konuları öğrenmeyi heyecanla beklerim.	1	2	3	4	5	6
48. Yabancıların ne düşündükleri üzerinde çalışmak anlamlıdır.	1	2	3	4	5	6
49. Sorgulayıcı oluşum en güçlü yanlarımdan birisidir.	1	2	3	4	5	6
50. Görüşlerimi destekleyecek olguları ararım, görüşlerime karşıt olanları değil.	1	2	3	4	5	6
51. Karmaşık problemleri çözmeye çalışmak eğlencelidir.	1	2	3	4	5	6

KATILMIYORUM (-)			KATILYORUM (+)			
1	2	3	4	5	6	
Hiç Katılmıyorum	Katılmıyorum	Kısmen Katılmıyorum	Kısmen Katılıyorum	Katılıyorum	Tamamen Katılıyorum	
52. Başkalarının düşüncelerini anlama yeteneğimden dolayı takdir edilirim.	1	2	3	4	5	6
53. Benzetmeler ancak karada yüzücü paleti ile yürümek kadar kullanışlıdır.	1	2	3	4	5	6
54. Beni mantıklı biri olarak tanımlayabilirsiniz.	1	2	3	4	5	6
55. Herhangi bir şeyin nasıl çalıştığını anlamak bana büyük zevk verir.	1	2	3	4	5	6
56. İşler zorlaştığında, arkadaşlarım sorunu çözmek için bana gelirler.	1	2	3	4	5	6
57. Bir problemle karşılaştığında, ilk yapılması gereken şey problemin ne olduğunu iyice anlamaya çalışmaktır.	1	2	3	4	5	6
58. Tartışmalı konulardaki fikrim genellikle en son konuştuğum kişiye bağlıdır.	1	2	3	4	5	6
59. Konu ne hakkında olursa olsun, daha fazlasını öğrenmek için can atarım.	1	2	3	4	5	6
60. Bir çözümün diğerinden daha iyi olup olmadığını bilmenin hiç bir yolu yoktur.	1	2	3	4	5	6
61. Soruları çözmeyen en iyi yolu, yanıtları başkasından istemektir.	1	2	3	4	5	6
62. Bir çok soru sorulamayacak kadar ürkütücüdür.	1	2	3	4	5	6
63. Karmaşık problemlere karşı düzenli yaklaşımımla tanınırım.	1	2	3	4	5	6
64. Farklı dünya görüşlerine karşı açık fikirli olmak, insanların düşündüğü kadar önemli bir şey değildir.	1	2	3	4	5	6
65. Öğrenebileceğin her şeyi öğren, ne zaman işe yarayacağını bilemezsin.	1	2	3	4	5	6
66. Hayat bana aşırı mantıklı olmamak gerektiğini öğretti.	1	2	3	4	5	6
67. Her şey görüldüğü gibidir.	1	2	3	4	5	6
68. Bir problem üzerinde uğraşmam gerektiğinde, diğer şeyleri kafamdan tamamıyla çıkartabilirim.	1	2	3	4	5	6
69. Diğer insanlar, sorunun çözümlenmiş sayılacağı kararını alırken bana danışırlar.	1	2	3	4	5	6
70. Kafamda bir fikir oluşmuşsa, seçenekleri değerlendiriyor gibi davranmama gerek yoktur.	1	2	3	4	5	6
71. Güçlü kişiler doğru cevabı kolayca belirleyebilirler.	1	2	3	4	5	6

KATILMIYORUM (-)			KATILIYORUM (+)		
1	2	3	4	5	6
Hiç Katılmıyorum	Katılmıyorum	Kısmen Katılmıyorum	Kısmen Katılıyorum	Katılıyorum	Tamamen Katılıyorum

72. Pek çok soruya hangi standartları uygulayacağımızı bilmek imkansızdır.	1	2	3	4	5	6
73. Başkaları her ne kadar kendi fikirlerini ortaya koysalarda benim onları duymaya ihtiyacım yoktur.	1	2	3	4	5	6
74. Karmaşık problemlerin çözümüne yönelik düzenli planlar geliştirmede iyiyimdir.	1	2	3	4	5	6
75. İnsanların benimle aynı fikirde olmalarını sağlamak için işe yarayabilecek her türlü sebebi gösterebilirim.	1	2	3	4	5	6

Appendix C: Means, Standard Deviations, and Pearson Moment Product Correlations across Turkish and US Samples

		TURKISH													
		X5	X9	X17	X22	X25	X27	X28	X32	X33	X35	X39			
	M	2.91	4.48	3.96	4	3.81	4.08	4.27	4.41	3.86	3.28	3.27			
	SD	1.592	1.514	1.590	1.369	1.650	1.283	1.666	1.570	1.660	1.547	1.451			
	M	3.25	1.571	1	-0.187	-0.084	-0.076	-0.060	-0.159	-0.045	-0.108	0.536	0.307		
	SD	4.13	1.558	-0.203	1	0.485	0.529	0.499	0.064	0.093	0.005	0.462	-0.131	-0.125	
	M	3.71	1.535	-0.086	0.595	1	0.372	0.368	0.042	0.109	0.015	0.423	-0.059	-0.067	
	SD	3.81	1.333	-0.156	0.589	0.464	1	0.435	0.080	0.167	0.174	0.404	0.023	-0.017	
US	M	3.54	1.588	-0.088	0.549	0.409	0.511	1	0.143	0.135	0.171	0.394	-0.096	-0.124	
	SD	3.82	1.210	-0.169	0.004	-0.071	0.052	0.054	1	0.261	0.318	0.037	-0.048	0.029	
	M	4.04	1.634	-0.135	0.033	0.072	0.195	0.186	0.282	1	0.431	0.130	-0.068	0.010	
	SD	4.19	1.506	-0.187	-0.022	0.043	0.100	0.134	0.373	0.539	1	0.084	-0.116	0.018	

Note: Correlations from U.S. sample (n=448) are shown below the diagonal; correlations from Turkish sample (n=583) are shown above the diagonal

Appendix C: Means, Standard Deviations, and Pearson Moment Product Correlations across Turkish and US Samples

		TURKISH											
		X5	X9	X17	X22	X25	X27	X28	X32	X33	X35	X39	
M		2.91	4.48	3.96	4	3.81	4.08	4.27	4.41	3.86	3.28	3.27	
SD		1.592	1.514	1.590	1.369	1.650	1.283	1.666	1.570	1.660	1.547	1.451	
M	SD												
X33	3.63	1.612	-0.012	0.462	0.423	0.404	0.394	0.037	0.130	0.154	1	0.082	0.087
X35	3.41	1.475	0.498	-0.156	-0.059	0.023	-0.096	-0.048	-0.068	-0.092	0.082	1	0.342
X39	3.33	1.332	0.381	-0.147	-0.111	-0.017	-0.124	0.029	0.010	0.046	0.087	0.342	1
X40	4.24	1.368	-0.040	0.013	-0.064	0.113	0.001	0.504	0.135	0.289	-0.094	0.044	0.029
X43	3.18	1.668	0.505	-0.175	-0.029	-0.094	-0.098	-0.088	-0.121	-0.018	0.146	0.323	0.446
X46	3.80	1.287	0.011	-0.191	-0.190	-0.041	-0.053	0.524	0.168	0.302	-0.054	-0.034	0.032
X49	4.08	1.441	-0.190	-0.062	-0.036	0.094	-0.077	0.388	0.212	0.267	-0.050	-0.025	-0.085
X52	4.10	1.374	-0.014	-0.046	-0.019	0.091	-0.005	0.301	0.213	0.268	-0.009	0.048	0.046

Note: Correlations from U.S. sample (n=448) are shown below the diagonal; correlations from Turkish sample (n=583) are shown above the diagonal

Appendix C: Means, Standard Deviations, and Pearson Moment Product Correlations across Turkish and US Samples

		TURKISH													
		X5	X9	X17	X22	X25	X27	X28	X32	X33	X35	X39			
	M	2.91	4.48	3.96	4	3.81	4.08	4.27	4.41	3.86	3.28	3.27			
	SD	1.592	1.514	1.590	1.369	1.650	1.283	1.666	1.570	1.660	1.547	1.451			
	M	3.96	3.96	3.96	3.96	3.96	3.96	3.96	3.96	3.96	3.96	3.96	3.96	3.96	
	SD	1.646	1.646	1.646	1.646	1.646	1.646	1.646	1.646	1.646	1.646	1.646	1.646	1.646	
X61		-0.106	0.051	0.056	0.120	0.117	0.436	0.477	0.510	0.110	-0.047	0.014			
X67		-0.107	-0.136	0.048	0.037	0.118	0.284	0.458	0.392	0.014	-0.111	-0.065			
X70		0.477	-0.229	-0.100	-0.128	-0.020	-0.137	-0.093	-0.032	0.000	0.255	0.318			
X72		0.527	-0.065	-0.047	-0.107	-0.046	-0.110	0.031	-0.024	0.109	0.393	0.388			
X74		-0.032	0.704	0.563	0.602	0.527	0.095	0.089	-0.023	0.453	-0.002	-0.004			
X75		0.367	0.001	0.135	0.011	0.052	-0.132	-0.050	0.001	0.189	0.311	0.362			

Note: Correlations from U.S. sample (n=448) are shown below the diagonal; correlations from Turkish sample (n=583) are shown above the diagonal

Appendix C: Means, Standard Deviations, and Pearson Moment Product Correlations across Turkish and US Samples

		TURKISH													
		X40	X43	X46	X49	X52	X61	X67	X70	X72	X74	X75			
M		4.43	2.93	3.88	4.32	4.51	4.60	4.28	3.17	3.60	3.87	3.21			
	SD	1.294	1.698	1.292	1.354	1.168	1.527	1.764	1.449	1.365	1.514	1.459			
US	M	3.25	1.571	3.71	1.535	3.81	1.333	3.54	1.588	3.82	1.210	4.04	4.19	1.506	
	SD	1.571	1.558	1.535	1.588	1.333	1.588	1.210	1.506	1.571	1.558	1.535	1.588	1.333	
	X5	-0.020	0.358	-0.011	-0.130	-0.038	-0.203	-0.163	0.352	0.410	-0.048	0.290			
	X9	0.013	-0.161	-0.072	0.035	0.049	0.051	-0.008	-0.169	-0.079	0.676	0.003			
	X17	-0.064	-0.029	-0.109	-0.013	-0.042	0.056	0.048	-0.083	-0.046	0.499	0.102			
	X22	0.091	-0.094	-0.041	0.148	0.150	0.120	0.037	-0.128	-0.034	0.582	-0.041			
	X25	-0.042	-0.084	-0.053	-0.077	0.057	0.117	0.118	-0.020	-0.046	0.484	0.023			
	X27	0.500	-0.079	0.497	0.388	0.301	0.346	0.284	-0.137	-0.110	0.095	-0.168			
	X28	0.195	-0.125	0.189	0.140	0.213	0.381	0.488	-0.093	0.031	0.089	-0.027			
	X32	0.328	-0.045	0.277	0.275	0.243	0.449	0.448	-0.039	-0.030	-0.112	0.000			

Note: Correlations from U.S. sample (n=448) are shown below the diagonal; correlations from Turkish sample (n=583) are shown above the diagonal

Appendix C: Means, Standard Deviations, and Pearson Moment Product Correlations across Turkish and US Samples

		TURKISH													
		X40	X43	X46	X49	X52	X61	X67	X70	X72	X74	X75			
M		4.43	2.93	3.88	4.32	4.51	4.60	4.28	3.17	3.60	3.87	3.21			
SD		1.294	1.698	1.292	1.354	1.168	1.527	1.764	1.449	1.365	1.514	1.459			
M	SD														
X33	3.63	1.612	-0.104	0.146	-0.054	0.025	0.110	0.014	0.000	0.109	0.453	0.145			
X35	3.41	1.475	0.034	0.403	-0.034	-0.025	0.034	-0.111	0.255	0.393	-0.002	0.294			
X39	3.33	1.332	0.053	0.466	-0.019	-0.085	0.023	0.014	-0.065	0.318	-0.004	0.334			
X40	4.24	1.368	1	0.000	0.548	0.540	0.479	0.371	0.199	-0.066	0.032	0.046	-0.082		
X43	3.18	1.668	0.000	1	0.018	-0.145	0.022	-0.123	-0.097	0.435	0.405	-0.112	0.403		
X46	3.80	1.287	0.548	0.018	1	0.520	0.445	0.360	0.243	-0.079	-0.043	-0.032	-0.116		
X49	4.08	1.441	0.570	-0.145	0.520	1	0.397	0.322	0.172	-0.157	-0.119	0.043	-0.177		
X52	4.10	1.374	0.392	0.051	0.445	0.397	1	0.405	0.175	0.062	0.096	0.069	-0.107		
US															

Note: Correlations from U.S. sample (n=448) are shown below the diagonal; correlations from Turkish sample (n=583) are shown above the diagonal

Appendix C: Means, Standard Deviations, and Pearson Moment Product Correlations across Turkish and US Samples

		TURKISH															
		X40	X43	X46	X49	X52	X61	X67	X70	X72	X74	X75					
	M	4.43	2.93	3.88	4.32	4.51	4.60	4.28	3.17	3.60	3.87	3.21					
	SD	1.294	1.698	1.292	1.354	1.168	1.527	1.764	1.449	1.365	1.514	1.459					
	M	SD															
X61	3.96	1.646	0.371	-0.123	0.398	0.382	0.386	1	0.495	-0.157	-0.002	0.084	-0.030				
X67	3.73	1.812	0.216	-0.012	0.243	0.234	0.295	0.495	1	-0.052	-0.070	-0.001	-0.034				
X70	3.25	1.430	-0.098	0.556	-0.004	-0.157	-0.018	-0.157	-0.052	1	0.398	-0.115	0.352				
X72	3.67	1.332	0.044	0.443	0.004	-0.116	0.095	-0.002	-0.070	0.398	1	0.008	0.309				
X74	3.64	1.555	-0.024	-0.089	-0.192	-0.060	-0.038	0.084	-0.001	-0.115	0.008	1	-0.003				
X75	3.35	1.428	-0.023	0.461	-0.093	-0.167	-0.017	-0.018	-0.034	0.352	0.309	-0.003	1				

Note: Correlations from U.S. sample (n=448) are shown below the diagonal; correlations from Turkish sample (n=583) are shown above the diagonal

Appendix C: Means, Standard Deviations, and Pearson Moment Product Correlations across Turkish and US Samples

		TURKISH													
		X40	X43	X46	X49	X52	X61	X67	X70	X72	X74	X75			
M	4.43	2.93	3.88	4.32	4.51	4.60	4.28	3.17	3.60	3.87	3.21				
SD	1.294	1.698	1.292	1.354	1.168	1.527	1.764	1.449	1.365	1.514	1.459				
M	SD														
X61	3.96	1.646	0.371	-0.12	0.398	0.382	0.386	1	0.495	-0.16	-0	0.084	-0.03		
X67	3.73	1.812	0.216	-0.01	0.243	0.234	0.295	0.495	1	-0.05	-0.07	-0	-0.03		
X70	3.25	1.430	-0.1	0.556	-0	-0.16	-0.02	-0.16	-0.05	1	0.398	-0.12	0.352		
US	X72	3.67	1.332	0.044	0.443	0.004	-0.116	0.095	-0	-0.07	0.398	1	0.008	0.309	
X74	3.64	1.555	-0.02	-0.09	-0.19	-0.06	-0.04	0.084	-0	-0.12	0.008	1	-0		
X75	3.35	1.428	-0.02	0.461	-0.09	-0.17	-0.02	-0.02	-0.03	0.352	0.309	-0	1		

Note: Correlations from U.S. sample (n=448) are shown below the diagonal; correlations from Turkish sample (n=583) are shown above the diagonal