Adaptation of Teachers' Self-Efficacy Towards Teaching Thinking Skills Scale Into English

Yal çın Dilekli[1], Erdoğan Tezci[2]

[1]Education Faculty, Aksaray University, Aksaray, Turkey

[2]Necatibey Eduaction Faculty, Balıkesir University, Balıkesir, Turkey

Correspondence: Yaçın DİLEKLİ, Education Faculty, Aksaray University, Adana Yolu 9. Km. Merkez/Aksaray, Turkey.

Abstract

Teaching thinking skills is core of the curriculum in many cultures. Although curricula have mutual points, the results were different from each other. Having different results of teaching thinking curriculums applied in different countries indicates importance of personal differences of teachers. Defining level of teacher's self-efficacy makes a contribution to determining success level of curriculum. Developing a global scale contributes both researchers and practitioner while teaching thinking. The aim of this study is to adapt Teachers' Self-efficacy towards Teaching Thinking Skills Scale (TSTS), developed in Turkish, into English. The scale consists of 20 items. Each item has 5 point Likert type. It has 3 factors as Design, Practice and Academic Competence. For this purpose, firstly linguistic equivalence was analyzed. For linguistic equivalence, both versions of the scale were applied to 28 candidate teachers of English with 20 day intervals. In the second step, Confirmatory Factor Analyses was administrated to data taken from 144 native English teachers. After the analysis, good level fitting indices were found. Cronbach Alpha coefficient value is .94. Construct validity (Convergence and Discriminating validity) study on correlations between sub-dimensions and average explained variance value has good sufficient level. Scale items were found discriminating. Results showed that English version of the scale was statistically valid and reliable.

Keywords: teaching thinking, self-efficacy, linguistic validity, scale adaptation

1. Introduction

'Even in our century, students don't know how to think' (McGrane & Sternberg, 1992: 339). Despite many efforts on curriculum development in order to grow up new thinking generation, we aren't at the desired level. If many countries applied similar programs, why obtained results are different? So, other variables on teaching thinking such as learners' and teachers' individual differences should be reconsidered again. Teachers' self-efficacy level which affect nearly teachers' all attitudes in classroom is one of them (Bandura & Schunk, 1981; Zimmerman & Kitsantas, 2005). Assessing teachers' level of self-efficacy becomes one of the important part in understanding causes of the different results.

2. Literature Review

In the last decade, there has been a new tendency in curriculum and instruction which is called thinking skills (Dilekli & Tezci, 2015; Zohar, 2013). However, definition or the scope of this term was not definitely defined up to 2001 (Costa, 2001; McGuinness, 1999; McGregor, 2007; Nispet, 1990; Wilks, 2005; Hashim, 2004; Tebbs, 2000; Alnesyan, 2012). 60 scientists defined scope of the term as problem solving, critical thinking, creative thinking and decision making (Costa, 2001). Definition of the term, the most difficult part of the problem, was solved. The second step was to develop or shape new curricula for teaching thinking. For this purpose, three approaches were applied in education. The first one was subject/content free programs for teaching thinking (Eg; Instrumental Enrichment by Fuerstein, Teaching Thinking Program by Sommerset, Top Ten Tactics by Lake and Needham, CoRT by De Bono) (Dilekli, 2015). Content free approach was criticized because of neglecting the real value of the knowledge or disciplines. An alternative approach was aroused which is based on content (Eg: Cognitive Acceleration Program, Thinking through Series by Adey and Sayer; Philosophy for Children Program, by Lipman) (McGregor, 2007). However, this approach was also criticized because of the problems occurred during transfer. Transferring thinking skills from one discipline to another was very difficult for learners, and they couldn't transfer this skill with the desired level of competency (Winch, 2010). The last approach launched by Swarts and Parks (2004:9) was called as infused approach. In this approach, components of

thinking skills were defined as key competencies of all the teaching curricula. In this approach, contents or subjects are used as tools for teaching thinking. At this point, educational activities done in the classroom become one of the most important part of teaching thinking (Dilekli &Tezci, 2015).

Apart from these approaches, some other programs were applied in many different countries for teaching thinking. The first program was *Odessay Program* in Venezuela and the second one was *Smart School Project* in America. For understanding the different results of the programs, other variables such as school atmosphere, teachers' attitudes, teaching techniques, teacher motivation, course books or course materials, students' individual differences, teaching and learning styles and teachers' self-efficacy level which affect on teaching thinking were hold in many researches (Alnesyan, 2012; Burke & Williams, 2008; Mansour, 2009; Zhang, 2006, Fisher, 2005; Newmann,1995; Baumfield, 2006, Dilekli & Tezci, 2015; Goelz,2004; Ritchhart & Perkins, 2000; McGrane & Sternberg, 1992; Wassermann, 2010; Dori & Hertzovich, 1999; Newmann & Wehlage, 2002; Tebbs, 2000; Zohar & Schwartzer, 2005, Dilekli & Tezci, 2016; Yenipınar & Akgün, 2017). Self-efficacy is one of the dominant figures on teaching styles while applying teaching techniques, teachers' motivation and creating friendly atmosphere for teaching thinking classrooms and schools (Dilekli & Tezci, 2016; Dilekli & Tezci, 2015; Tebbs, 2010). In many study (Baumfield & Oberski, 1998; Nair & Ngang, 2012, Wasserman, 2010; Tebbs, 2010; Dilekli & Tezci, 2015; Soh, 2017), classroom climate was mentioned as one of the important barriers standing in front of teaching thinking process. Therefore, classroom climate has strictly connected with teacher's self-efficacy level (Alnesyan, 2012; Alwehaibi, 2012; Beyer & Ronald, 1986; Dilekli, 2015; Otman & Muhammad, 2014).

The term self-efficacy is *defined as people's beliefs about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives* (Bandura, 1994:72*).* Self-efficacy is effective on our feelings, way of thinking and motivation (Zimmerman & Bandura & Martinez-Ponz, 1992). Self-efficacy makes learners and teachers more enthusiastic on their jobs. As teaching thinking is a long process, teachers and learners should be patient to reach the desired level. Teachers having low self-efficacy levels give up teaching thinking when they face a problem (Onosko, 1991; Hashim, 2004; Alnesyan, 2012). Similarly, Ashton and Webb (1986) indicated that teachers with low self-efficacy were easily discouraged while teaching the subjects needing long time and effort. However, teachers having high self-efficacy level were more successful in teaching thinking skills and their students' level of self-efficacy increased meaningfully (Hampton, 1996). Measuring teachers' self-efficacy level allows us to understand teachers' attitudes and behaviors (Kaya, 2008). Assessing teachers' self-efficacy level helps us to design better curriculum for teaching thinking. However, assessing self-efficacy is not easy, so it is difficult to mention about a global instrument assessing self-efficacy. Notwithstanding, there are many other self-efficacy culture specific scales for different disciplines or specific subject-matters.

2. Method

2.1 Participants

In this study, it was aimed to analyze linguistic validity and reliability of TSTS scale's English version. For this purpose, the translation of the scale was checked by two native English Language experts knowing Turkish and two native Turkish Language experts knowing English. These experts were academic staff of Turkish and English languages teaching departments with Ph.D. degree. Furthermore, in order to check linguistic equivalence of the translated version of the scale, in the first step, it was applied to 28 candidate teachers of English face to face with 20-day interval. 10 of these candidate teachers were male and 18 of them were female. In the second step, the data were collected from 144 volunteer teachers whose native language is English for validity and reliability analysis of TSTS scale. 112 (77.8 %) of them were female and 32 (22.2%) of them were male. Participants' professional seniority were as follows; 8 (5.6%) of them had 0-5 year experience, 22 (15.3%) of them had 6-10 year experience, 24 (16.7%) them had 11-15 year experience, 27 (18.8%) of them had 16-20 year experience and 63 (43.8%) of them had 21 and over year experience. The data were collected via internet by sending e-mails directing to the survey link. Besides, the teachers having this link were free to share other volunteer teachers.

2.2 Data Collection Tool

Turkish version of TSTS scale consisted of 20 items. The scale is 5 point Likert scale, 5 = Strongly Agree, 4= Agree, 3 = Undecided, 2= Disagree and 1= Strongly Disagree. After Explanatory Factor Analysis (EFA), it was found that the scale developed by Dilekli and Tezci (2015) has three factors. The first factor called Academic Competence (Competence) explains 27,878% of the variance and consisted of 8 items, the second factor Practice explains 22.637% variance and consisted of 8 items, the last factor Design explains 12.259% of the variance and consisted of 4 items. The total explained variance is 62.774%. Cronbach Alpha Coefficient values were as follows; for Academic Competence (Competence) .89, for Practice .93 and for design .74; for overall the scale, it was found .95.

For linguistic equivalence, correlational analysis and paired samples t-test were administrated to the data collected form

28 candidate teachers of English. Then confirmatory factor analysis (CFA) were administrated to the data collected via emails from 144 teachers whose native language is English. CFA is a statistical technique used to verify the factor structure of a set of observed variables or their underlying latent constructs exists latent variables (Büyüköztürk, Şekercioğlu & Çokluk, 2014). CFA analysis was used to test the relationships among factors and which variables related to which factors, the level of relationships among the factors and whether the data fit a hypothesized measurement model (Bentler & Bonett, 1980; Tabachnick & Fidell, 2007). In other words, CFA was used to test meaningfulness of the relationships between the observed variables and the theoretical structure defined by Explanatory Factor Analysis (EFA).

For defining fitness of the factor structure, different indices can be used. Firstly, Chi-Square and p value were checked. Later, AGFI (Adjusted Goodness of Fit Index) and GFI (Goodness Fit Index) values were analyzed. The value 1 indicates perfect fit for AGFI and GFI. Furthermore, RMSEA (Root Mean Square Error of Approximation), RMR (the Root Mean Residual) and SRMR (the Standardized Root Mean Square Residual) values were calculated. PGFI (Parsimony Goodness-of-Fit Index) is another indices showing the simplicity. Another indices CFI (Comparative Fit Index) is used for small sampling groups. NFI (The Normed Fit Index) and NNFI (Non-Normed Fit Index) indices let researchers compare and analyze the results without required scattering of the Chi-Square. PGFI (Parsimony Goodness-of-Fit Index) is another indices showing plainness of the scale (Bollen 1989; Hoyle 1995; Jöreskog & Sörbom, 1996; Tabachnick & Fidell, 2007). In order to convergent validity, Average Variance Extracted (AVE) and Composite Reliability (CR) indices were checked. AVE indices should be higher than .50, but if CR index is higher than .60, having higher than .40 value for AVE index is acceptable (Fornell & Larcker, 1981). In order to see internal consistency of the adopted scale, Cronbah's Alpha analysis was administrated. For defining item discriminating indices, top and low group 27% technique was used. Furthermore, for item discriminating of a dimension (Practice, Design and Competence) and items independent sample t test were administrated between the groups formed by means of the 27% technique.

3. Results

3.1 Translation Process and Linguistic Equivalence

Firstly, the scale was translated by the researchers. Later, three experts, teaching English, whose mother tongue is Turkish checked the translation and in this step no correction was made. Then, three experts, teaching English, whose mother tongue is English and checked the translation, again no correction was made at this stage. After forming English version of the scale, both versions of the scale were applied to 28 candidate teachers of English at 20 day intervals. T-test and correlational analysis results of these applications were given in Table 1.

Table 1. Correlation and t Test Analysis Results of Turkish and English Versions of the Scale

		Mean	SD	Correlation		t	p
				r	P		
Item1	Turkish	4.04	.79	.694	.000*	.328	.745**
	English	4.00	.47				
Item2	Turkish	3.96	.79	.702	.000*	-1.000	.326**
	English	4.07	.60				
Item3	Turkish	3.75	.75	.748	.000*	.372	.713**
	English	3.71	.66				
Item4	Turkish	4.18	.77	.804	.000*	1.140	.264**
	English	4.07	.81				
Item5	Turkish	3.71	.94	.781	.000*	.297	.769**
	English	3.68	.98				
Item6	Turkish	3.79	1.11	.852	.000*	.328	.745**
	English	3.75	.97				
Item7	Turkish	4.00	.72	.574	.001*	-.570	.573**
	English	4.07	.72				
Item8	Turkish	4.04	.79	.690	.000*	-.626	.537**
	English	4.11	.74				
Item9	Turkish	4.18	.67	.744	.000*	.812	.424**
	English	4.11	.63				
Item10	Turkish	4.29	.85	.490	.008*	.721	.477**
	English	4.18	.67				
Item11	Turkish	3.61	.99	.860	.000*	.372	.713**
	English	3.57	.88				
Item12	Turkish	4.14	.59	.794	.000*	.812	.424**
	English	4.07	.77				
Item13	Turkish	3.54	1.10	.862	.000*	-1.000	.326**
	English	3.64	.87				
Item14	Turkish	3.50	.92	.757	.000*	-.626	.537**
	English	3.57	.74				
Item15	Turkish	3.68	.90	.839	.000*	-1.140	.254**
	English	3.79	.69				
Item16	Turkish	3.46	1.03	.751	.000*	.827	.415**
	English	3.36	.83				
Item17	Turkish	3.64	.99	.768	.000*	.297	.769**
	English	3.61	.83				
Item18	Turkish	3.79	.88	.729	.000*	.626	.537**
	English	3.71	.71				
Item19	Turkish	3.61	1.03	.825	.000*	-1.279	.212**
	English	3.75	.75				
Item20	Turkish	3.57	.99	.784	.000*	-.902	.375**
	English	3.68	.67				

Df=27, *p<.05; **p>.05

Analysis results showed that there were positive and meaningful relationships among the scale items. The lowest correlation (r=.574, p<.05) was seen in item 7 and the highest correlation (r=.862, p<.05) was seen in item 13. The results obtained from both versions of scale showed medium and perfect degree of correlation. However, according to the t test results, both versions of the scale items didn't have meaningful relationship (p>.05).

3.2 Results of CFA Analysis

CFA analysis was administrated in order to determine whether the factor structure of English version, which consists of three factors, was the same its Turkish version or not. According to the results of CFA analysis, the value of CFA was found as X^2/df =305.88/167=1.83 and the RMSEA value in this study was found =.76. These values are between acceptable levels. However, some other values showing the simplicity of the scale were found low (GFI=.88; AGFI=.78). After the proposed three modifications (between item 19 and 20; item 5 and 9; 13 and 14), better fitting indices were seen. Analysis results were given in Table 2.

Table 2. Goodness-of-fit indices for CFA according to three factor model

	X^2	Df	RMSEA	NFI	NNFI	CFI	IFI	RFI	RMR	SRMR	GFI	AGFI
DFA	305.88	167	.076	.95	.97	.98	.98	.94	.066	.058	.87	.82
3 Modification	249.30	164	.060	.96	.98	.98	.98	.95	.065	.057	.90	.85

After the proposed modifications based on the error variance were made, the fitting indices RMSEA value went down

from .076 to .060 and NFI indices increased from .95 to .96. These indices show excellent fitting levels. AGFI indices showing the simplicity of the model increased from .82 to .85. The path of coefficients, from observed variables to latent variables, were meaningful. The standardized path diagram was given in Figure 1.
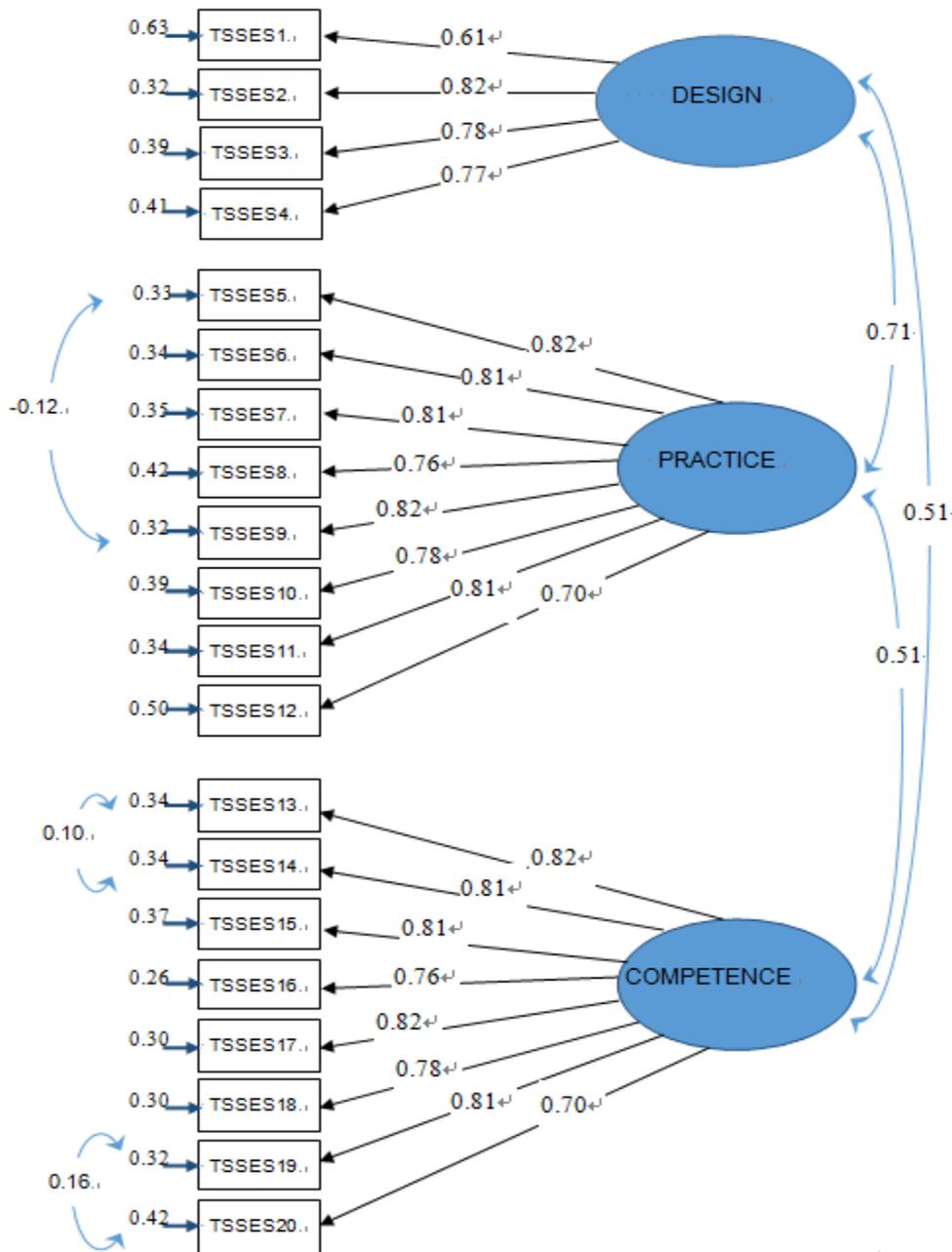

Figure 1. Parameters of Standardized CFA

As it was seen in the Figure 1, in Design dimension, the highest path coefficient was seen .82 (t=11.38, p<.05) in item 2 and the lowest path coefficient was .61 (t=7.53, p<.05) in item 1. For Practice dimension, the highest path coefficient was .82 (t=11.83 and t=11.73, p<.05) in items 5 and 9, the lowest path coefficient was .70 (t=9.51, p<.05) in item 12. For Competence dimension, the highest path coefficient was .86 (t= 12.79, p<.05) in item 16 and the lowest one .76 (t=10.56, p<.05) in item 20.

3.3 Convergent and Discriminant Validity

Average Variance Extracted (AVE) values and Composite Reliability (CR) scores were shown in Table 3. All AVE values were higher than .50, which provide evidence for the convergent validity of our scale. All CR values were expected to be higher than .70 (Gouveia & Soares, 2015).

Table 3. AVE, CR Values and Correlations among Dimensions

	AVE	CR	1	2	3
1-Design	.56	.84	(.75)		
2-Practice	.63	.93	.71	(.79)	
3-Competence	.67	.94	.52	.51	(.82)

Note: Square root of average variances extracted are shown on diagonal.

Having over .50 factor load and AVE values show the convergent validity of the scale (Fornel & Larcker, 1981; Peterson, 2000). The lowest factor load was .61 in item 1 and the highest factor load .86 in item 16. The rest were between these values. As it was seen in the Table 3. AVE values were found for Design dimension .56; for Practice dimension .63 and for Competence dimension .67.

3.4 Reliability and Discrimination Indices

Cronbach Alpha and Omega reliability were calculated for reliability and internal validity. Furthermore, item total variances correlations and item discriminations indices were calculated. The results were shown in Table 4.

Table 4. Reliability and Item Total Correlations

	Mean	SD	Skewness	Kurtosis	Alpha	Omega Reliability(ω)
1-Design	3.72	.95	-.9454	.3631	.85	.85
2-Practice	3.76	.86	-.9533	.7965	.94	.94
3-Competence	3.72	.87	-.5977	-.1678	.95	.95
General	3.73	.73	-.5194	.7652	.94	.94

Cronbach's alpha values providing evidence for the reliability of the scale were found .85 for Design, .94 for Practice, and .95 for Competence factors. Omega reliability indices of all sub factors are the same Cronbach's Alpha values. These values indicated that the participants replied in a consistent manner to the scale items. Corrected item total correlation scores ranged from .49 to .63 for Design, .58 to .75 for Practice, .60 to .71 for Competence dimensions. These values indicated that all of the items on the other dimensions were moderately correlated with total scores on their respective dimension. The t values for the top 27 and lowest 27 percent of the students were significant and ranged from 5.669 to 19.411, which provides evidence supporting the discriminatory power of the scale items.

4. Discussion

The aim of this study is to adapt into English TSTS scale developed in Turkish. Firstly, the scale translated into English and the translated scale applied to 28 candidate teachers of English at 20 day intervals in order to analyze its understandability. Paired samples t test and correlation analysis were administrated with the data taken from this application. It was found medium and high correlation between Turkish and English versions of the scale, but it was not found meaningful relationship between the two versions of the scale in paired samples t test analysis. As a result, linguistic equivalence analysis depicted that English version of the scale is understandable and clear and refers to similar meanings with Turkish version.

In the second step of the adaptation process, translated scale was applied to 144 native English teachers. After the CFA administrated to 144 participants' data, it was seen that the scale has acceptable indices in respect for RMSEA, RMR, SRMR on the other hand some GFI and AGFI values were found low. Then three proposed corrections based on error variances were administrated. With these proposed modifications, GFI and AGFI values reached acceptable level. Such as; RMSEA went down from .076 to .060 and RMR from .66 to .65, SRMR from .58 to 57.GFI and AGFI vales found .90 and .85 which were between acceptable levels. Consequently, English version of TSTS scale have the same factor structure with its Turkish version. This finding is similar to Dilekli and Tezci (2015) who developed this scale. But, Dilekli and Tezci (2015) didn't make modifications based on error variances.

Convergent validity and discriminant validity analyses were administrated. All the factor load values were higher than .50 (the values range from .61-.86) and each dimension AVE indices were higher than .50, these values showed that the scale has convergent validity. According to Fornel and Larcker (1981), and Peterson (2000), having higher than .50 AVE value shows convergent validity of a scale. For discriminant validity, correlation analyses were administrated between square root of AVE values and relationships of each sub-dimensions of scale. The results indicated that English version of the scale has discriminant validity. Omega Reliability and Cronbach Alpha values were found high, so it could be said that English version of the scale has internal consistency. Top and low 27% technique was administrated for item discriminating values. The results showed that items were discriminating. The reliability co-efficient found in

this study is higher than the study by Dilekli and Tezci (2015). There are many teachers' self-efficacy scale special for some disciplines (Eg; Enochs & Riggs, 1990; Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998), but there are very limited number of scale for assessing teachers' self-efficacy for teaching thinking skills. One of these scale developed by Tebbs (2000) has 4 factors. The scale was adapted by Kaya (2008) into Turkish. However, Kaya found that the adopted version has a different construct from the original version of the scale. This shows that linguistic and cultural differences have an important role in adaptation process. Stes, De Maeyer & Van Petegem, (2010) indicated that translation and cultural differences are effective factors in the adaptation process. Similarly, Meyer & Eley (2006) draw attention that any scale shouldn't be used without adaptation of a scale. Schwarzer, Bäßler, Kwiatek, Schröder, & Zhang (1997) studied on adaptation of a scale, originally in German, into Spanish and Chinese culture. Although they found meaningful relationship between means and gender, they didn't get satisfactory results in reliability and validity analysis. However, Luszczynska, Scholz, & Schwarzer (2005) found that self-efficacy is a global construct in their meta- analysis study from different cultures. TSTS scale can be used to define level of self-efficacy of the teachers teaching any discipline as its items are prepared based on infused teaching thinking skill approach. Moreover, it is a self-report survey which makes easier to collect data. As TSTS items were prepared according to the infused approach principles, it gave opportunity to compare teacher self-efficacy levels teaching different disciplines.

Onosko (1991) found that teachers' self-efficacy is one of the main problem in teaching thinking skills. Having low self-efficacy results in selecting wrong or inappropriate teaching techniques in class (Hoy & Spero, 2005). Andersen, Dragsted, Evans, & Sorensen (2004) and Pearson & Moomaw (2005) found that teachers having low self-efficacy aren't ready to take risk or trying new techniques in the classroom. From this respect, giving education on teaching thinking during the preservice period in education faculty will increase their self-efficacy level. Klassen, Tze, Betts & Gordon (2011) concluded that as self-efficacy scales based on self-report measures, there may be inconsistency between the findings of the scale and teachers' practice in the classroom. In fact, the disagreements about self-efficacy scale arise from that the most scales are for special disciplines or teachers' personal characteristic effecting the measurement process. For this reason, measurements related to self- efficacy are more related to attitudes of the teachers towards practice than actual practice. The adopted scale in this study based on infusion approach principles explained by McGuinnes (1997) and Swarts and Parks (1994) as pedagogical principles of thinking skills.

Bandura (1997) indicated that self-efficacy scale should be focused on belief based on the current ability. But, some study on developing self-efficacy scales either measures abilities (Ho & Hau, 2004) or capacity to apply special abilities or skills. For this reason, during the adaptation process of a scale, translation of items is very important step as it may cause different semantic results. Therefore, the structural equivalence of the scale, both linguistically and culturally should be ensured that the translation is as accurate as possible (Orakcı, 2018). Klassen et all. (2009) adapted TSES scale developed by Tschannen-Moran & Woolfolk Hoy (2001). During the adaptation process, researchers collected data from 6 different countries and reached similar validity and reliability results. Moreover, they also found relationships between different constructs and self-efficacy of teachers from different countries. Values from different cultures and level of importance to the values, beliefs and cultural attitudes could affect educational values and practices. However, having similar results between adapted and original versions of the scale supports the idea that self-efficacy is a global construct (Luszczynska and et all., 2005). Similarly, the study by Scholz, Doña, Sud & Schwarzer (2002) reached similar results in their study, whose data were collected from 25 different countries, about teachers' self-efficacy.

5. Limitations

In this study, it is not analyzed for teaching critical thinking skills and criterion validity of the scale for defining perception of self-efficacy. It is also not analyzed relationship with different constructs (Eg; self-confidence, atttudes). Researches that handle relationship between the scale and different constructs make a significant contribution to global value of the scale.

References

Alnesyan, A. (2012). *Teaching and learning thinking skills in the kingdom of saudi arabia: A case studies from seven primary schools.* Unpublished PhD. Dissertation. University of Exeter, UK.

Alwehaibi, U. H. (2012). A Proposed program to develop teaching for thinking in pre-service english language teachers. *English Language Teaching, 5*(7), 53-63. https://doi.org/10.5539/elt.v5n7p53

Andersen, A. M., Dragsted, S., Evans, R. H., & Sorensen, H. (2004). The relationship between changes in teachers' self-efficacy beliefs and the science teaching environment of Danish first-year elementary teachers. *Journal of Science Teacher Education, 15*(1), 25-38. https://doi.org/10.1023/B:JSTE.0000031461.68912.3d

Ashton, P. T., & Webb, R. B. (1986). *Making a difference: teachers' sense of efficacy and student achievement.* New York: Longman.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and ıntrinsic ınterest through proximal self-motivation. *Journal of Personality and Social Psychology, 41,* 586-598. https://doi.org/10.1037/0022-3514.41.3.586

Baumfield, V., & Oberski, I. (1998). What do teachers think about thinking skills? *Quality Assurance in Education, 6*(1), 44–51. https://doi.org/10.1108/09684889810200386

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588-606. https://doi.org/10.1037/0033-2909.88.3.588

Beyer, B. K., & Ronald, E. (1986). Teaching thinking skills in biology. *The American Biology Teacher, 48*(4), 207-212. https://doi.org/10.2307/4448265

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research, 17*(3), 303-316. https://doi.org/10.1177/0049124189017003004

Burke, L. A., & Williams, J. M. (2008). Developing young thinkers: An intervention aimed to enhance children's thinking skills. *Thinking Skills and Creatitivity, 3*(2), 104-124. https://doi.org/10.1016/j.tsc.2008.01.001

Büyüköztürk, Ş., Çokluk, Ö. & Şekercioğlu, G. (2014). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve Lisrel uygulamaları.* Ankara: Pegem A yayıncılık.

Costa, L. A. (2001). *Developing Minds: A Resource Book for Teaching Thinking.* ASCD. Melbourne. AU.

Dilekli, Y. (2015). *The relationship among teachers' classroom practices for teaching thinking skills, teachers' self-efficacy towards teaching thinking skills and teachers' teaching styles.* Unpublished PhD. Dissertation, Institute of Social Science, Balıkesir University, Balıkesir/ Turkey.

Dilekli, Y., & Tezci, E. (2015). The study of validity and reliability of teachers' self-efficacy towards teaching thinking skills scale. *The Journal of Academic Social Science Studies.*Vol.38, 135-153. https://doi.org/10.9761/JASSS3056

Dilekli, Y., & Tezci, E. (2016) The relationship among teachers' classroom practices for teaching thinking skills, teachers' self-efficacy towards teaching thinking skills and teachers' teaching styles. *Thinking Skills and Creativity, Vol. 21*, *pp*.144–151. https://doi.org/10.1016/j.tsc.2016.06.001

Dori, Y. J., & Herscovitz, O. (1999). Question posing capability as an alternative evaluation method: Analysis of an environmental case study. *Journal of Research in Science Teaching, 36*(4), 411–430. https://doi.org/10.1002/(SICI)1098-2736(199904)36:4<411::AID-TEA2>3.0.CO;2-E

Enochs, L. G., & Riggs, I. M. (1990). Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *School Science and Mathematics, 90*(8), 694-706. https://doi.org/10.1111/j.1949-8594.1990.tb12048.x

Fisher, A. (2005). *Thinking skills and admission to higher education.* www.cambridgeassessment. org.uk /1139 Accesed: 05.10.2012.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 1*8(1), 39-50. https://doi.org/10.2307/3151312

Goelz, L. (2004). *The teacher beliefs and practice: Consistency or inconsistency in the high school social studies classroom? studies in teaching 2004: Readers Digest.* Winston- Salem USA.

Gouveia, V. V., & Soares, A. K. S. (2015). *Calculadora de validade de construto (CVC).* Joao Pessoa, PB: BNCS /Universidade Federal da Paraiba.

Guskey, T. R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education, 4*(1), 63-69. https://doi.org/10.1016/0742-051X(88)90025-X

Hashim, H. (2004). Malaysian teachers' attitudes, competency and practices in teaching thinking. *Intellectual Discourse, 11*(1), 27-50.

Ho, I. T., & Hau, K. T. (2004). Australian and Chinese teacher efficacy: Similarities and differences in personal instruction, discipline, guidance efficacy and beliefs in external determinants. *Teaching and Teacher Education, 20,* 313–323. https://doi.org/10.1016/j.tate.2003.09.009

Hoy, A. W., & Spero, R. B. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four

measures. *Teaching and Teacher Education, 21*(4), 343-356. https://doi.org/10.1016/j.tate.2005.01.007

Hoyle, R. H. (1995). *Structural equation modeling: Concepts, issues, and applications*. London: Sage Pub.

Johnson, S., & Siegel, H. (2010). *On thinking skills.* (Ed. Winch, C.), Teaching thinking: Key debates in educational policy, (113-125). London: UK Continuum Press.

Kaya, B. (2008). *Evaluation social science candidate teachers' self-efficacy level towards teaching thinking.* Unpublished PhD. Dissertation, Institute of Educational Science, Gazi University, Ankara/ Turkey. (Sosyal bilgiler öğretmen adaylarının düşünme becerilerinin öğretimine yönelik öz yeterliklerinin değerlendirilmesi. Yayınlanmamış Doktora Tezi, Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara).

Klassen, R. M., Bong, M., Usher, E. L., Chong, W. H., Huan, V. S., Wong, I. Y. F., & Georgiou, T. (2009). Exploring the validity of a teachers' self-efficacy scale in five countries. *Contemporary Educational Psychology, 3*4(1), 67–76. https://doi.org/10.1016/j.cedpsych.2008.08.001

Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998-2009: Signs of progress or unfulfilled promise? *Educational Psychology Review, 23*(1), 21-43. https://doi.org/10.1007/s10648-010-9141-8

Luszczynska, A., Scholz, U., & Schwarzer, R. (2005). The general self-efficacy scale: Multicultural validation studies. *The Journal of Psychology, 139*(5), 439-457. https://doi.org/10.3200/JRLP.139.5.439-457

Mansour, N. (2009). The experiences and personal religious beliefs of Egyptian science teachers as a framework for understanding the shaping and reshaping of their beliefs and practices about science-technology-society (STS). *International Journal of Science Education, 30*(12), 1605-1634. https://doi.org/10.1080/09500690701463303

McGrane, P. A., & Sternberg, R. J. (1992). Discussion: Fatal vision—The failure of the schools in teaching children to think. (Ed. C. Collins & J. N. Mangieri). *Teaching thinking: An agenda for the 21st century (333–344).* Hillsdale, NJ: Lawrence Erlbaum Associates.

McGregor, D. (2007). *Developing thinking: Developing learning.* New York, USA: Open University Press.

McGuinness, C. (1999). *From thinking skills to thinking classrooms.* DfEE Publications. Nottingham: UK. https://www.education.gov.uk/publications/ standard/ publicationDetail/Page1/RB115 Accessed: 05.10.2012.

Meyer, J. H. F., & Eley, M. G. (2006). The approaches to teaching inventory: A critique of its development and applicability. *British Journal of Educational Psychology, 76,* 633-649. https://doi.org/10.1348/000709905X49908

Nair, S., & Ngang, T. K. (2012). Exploring parents' and teachers' views of primary pupils' thinking skills and problem solving skills. *Creative Education, 3*(1), 30-36. https://doi.org/10.4236/ce.2012.31005

Newmann, F. M., & Wehlege, G. G. (1995). *Successful school reconstructing: A report to the publican educators.* Madison, WI: Wisconsin Center for Education Research, University of Wisconsin.

Nispet, D. J. (1990). *Teaching thinking: An introduction to the research literature. UK*: Scottish Council for Research in Education

Onosko, J. J. (1991). Barriers to the promotion of higher-order thinking in social studies. *Theory & Research in Social Education, 19*(4), 341-366. https://doi.org/10.1080/00933104.1991.10505646

Orakcı, Ş. (2018). İngilizce konuşma kaygısı ölçeği'nin geçerlik ve güvenirlik çalışması. *İnönü Üniversitesi Eğitim Bilimleri Enstitüsü Dergisi*, *5*(9), 1-28. https://doi.org/10.29129/inujgse.375916

Othman, N., & Mohamad, K. A. (2014). Thinking skill education and transformational progress in Malaysia. *International Education Studies, 7*(4), 22-34. https://doi.org/10.5539/ies.v7n4p27

Pearson, C. L., & Moomaw, W. (2005). The relationship between teacher autonomy and stress, work satisfaction, empowerment, and professionalism. *Education Research Quarterly, 20*(1), 37-53.

Ritchhart, R., & Perkins, N. D. (2004). *Learning the think: The challenges of teaching thinking.* (Eds: Hoyak, K. & R., G., Morrison). The Cambridge handbook of thinking and reasoning. (775-802). UK: Cambridge University Press.

Scholz, U., Doña, B. G., Sud, S., & Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment, 18*(3), 242-251. https://doi.org/10.1027//1015-5759.18.3.242

Schwarzer, R., Baßler, J., Kwiatek, P., Schröder, K., & Zhang, J. X. (1997). The assessment of optimistic self-beliefs: comparison of the German, Spanish, and Chinese versions of the general self-efficacy scale. *Applied Psychology, 46*(1), 69-88. https://doi.org/10.1111/j.1464-0597.1997.tb01096.x

Soh, K. (2017). Fostering student creativity through teacher behaviors. *Thinking Skills and Creativity, 23,* 58-66. https://doi.org/10.1016/j.tsc.2016.11.002

Stes, A., De Maeyer, S., & Van Petegem, P. (2010). Approaches to teaching in higher education: Validation of a Dutch version of the approaches to teaching inventory. *Learning Environments Research, 13*(1), 59-73. https://doi.org/10.1007/s10984-009-9066-7

Swartz, R., & Parks, S. (1994). *Infusing the teaching of critical and creative thinking into content instruction: A lesson design handbook for the elementary grades*. Pacific Grove, CA: Critical Thinking Books and Software.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Allyn & Bacon/Pearson Education.

Tebbs, J. (2000). *Assessing Teachers' Self-Efficacy towards Teaching Thinking.* PhD. Dissertation. University of Connecticut University, Connecticut.

Tschannen-Moran, M., & Woolfolk-Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17,* 783–805. https://doi.org/10.1016/S0742-051X(01)00036-1

Tschannen-Moran, M., Woolfolk-Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*(2), 202-248. https://doi.org/10.3102/00346543068002202

Wasserman, S. (2010). Teaching for thinking Redux. *Phi Delta Kappan, 91*(5), 81-84. https://doi.org/10.1177/003172171009100523

Wilks, S. (2005*). Designing a thinking curriculum*. Sidney: Acer Press.

Yenipınar, Ş., & Akgün, N. (2017). The level of implementation of the strategic management in primary schools, *Universal Journal of Educational Research 5*(5), 791-798. https://doi.org/10.13189/ujer.2017.050512

Zhang, L. (2006). Preferred teaching styles and modes of thinking among university students in Mainland China. *Thinking Skills and Creativity, 1,* 95–107. https://doi.org/10.1016/j.tsc.2006.06.002

Zimmerman, B. J., & Bandura, A., & Martinez-Pons, M. (1992) Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 2*9(3), 663-676. https://doi.org/10.3102/00028312029003663

Zimmerman, J. B., & Kitsantas, A. (2005). Homework practices and academic achievement: The mediating role of self-efficacy and perceived responsibility beliefs. *Contemporary Educational Psychology, 30,* 397–417. https://doi.org/10.1016/j.cedpsych.2005.05.003

Zohar, A. (2013). Challenges in wide scale implementation efforts to foster higher order thinking (HOT) in science education across a whole school system. *Thinking Skills and Creativity, 10,* 233-249. https://doi.org/10.1016/j.tsc.2013.06.002

Zohar, A., & Schwartzer, N. (2005). Assessing teachers' pedagogical knowledge in the context of teaching higher-order thinking. *International Journal of Science Education, 2*7(13), 1595-1620. https://doi.org/10.1080/09500690500186592

Appendix A

Teachers' Self-efficacy towards Teaching Thinking Skills Scale (TSTS)

	Scale Items	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1	I am able to find out which of the activities given in guide books (teachers' books) are related to teaching thinking skills.	1	2	3	4	5
2	I can plan/arrange a subject discussed in course books to teach thinking skills.	1	2	3	4	5
3	I can apply assessment techniques which make assessing thinking skills possible.	1	2	3	4	5
4	I can prepare classroom activities for teaching thinking skills.	1	2	3	4	5
5	I know what to do to improve thinking skills of the students who have difficulty on learning them.	1	2	3	4	5
6	I know how to design a reading text or a story in the course book to teach thinking skills.	1	2	3	4	5
7	I am able to cope with the problems which may occur during the process of teaching thinking skills.	1	2	3	4	5
8	I can design activities to improve self-learning abilities of students instead of teaching them.	1	2	3	4	5
9	During an experiment or observation, I can organize activities getting students to think.	1	2	3	4	5
10	I cite real life scenes and examples to teach thinking skills.	1	2	3	4	5
11	I know how to evaluate/assess improvement of thinking skills.	1	2	3	4	5
12	I know how to motivate students to think.	1	2	3	4	5
13	I know what are the components required during improving of thinking skills.	1	2	3	4	5
14	I can arrange activities intended to improve higher order thinking skills.	1	2	3	4	5
15	I know which activity improves which thinking skill.	1	2	3	4	5
16	I have adequate knowledge of thinking skills.	1	2	3	4	5
17	I feel competent to determine what thinking skills are.	1	2	3	4	5
18	I can design activities according to types of thinking (Eg: critical thinking, creative thinking etc.).	1	2	3	4	5
19	I have adequate experience on how to promote attainments of thinking skills given in curriculum.	1	2	3	4	5
20	I have adequate knowledge on how to promote attainments of thinking skills given in curriculum.	1	2	3	4	5