Developing and Validating a Scale to Measure Turkish and Kazakhstani ELT Pre-Service Teachers' Intercultural Communicative Competence

Lazura Kazykhankyzy
PhD Candidate, Hacettepe University, Turkey, *lazura.kazykhankyzy@gmail.com*

Nuray Alagözlü
Assoc. Prof., Hacettepe University, Turkey, *nurayalagozlu@gmail.com*

The purpose of this study is to describe the development of the Intercultural Communicative Competence Scale (ICCS). Two different groups from Turkey and Kazakhstan constituted the sample. The Turkish group (N=258) participated in the piloting process to develop the scale. The Kazakhstani group (N=314) was used to confirm the structure validity of the instrument. Based on review of literature, 76 items important for assessing ICC were generated. After expert revision for the face and content validity of the scale, the exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were applied. According to the results, the content validity index (CVI) for the ICC scale was found acceptable (CVI=0.7). EFA results found that the scale, consisting of 4 factors and 52 items, explained 50.78% of the total variance. The factor loadings showed that 21 items loaded fairly strongly on Component 1 (Skills), 13 items on Component 2 (Attitude), 12 items on Component 3 (Awareness), and 6 items on Component 4 (Knowledge). CFA results also confirmed the four factorial structure of the scale. The Cronbach's alpha reliability coefficients of four factors and a whole instrument were found .946; .906; .880; .806 and .958 respectively.

Keywords: intercultural communicative competence, ELT pre-service teachers, scale, teaching, learning

INTRODUCTION

The importance of Intercultural Communicative Competence (ICC) in foreign language learning and teaching has been attended to worldwide for many years. It was unanimously accepted by researchers and educators that ICC helps learners understand other cultures and develop global perspectives. Since the "ability to communicate effectively and appropriately in intercultural situations is based on one's intercultural knowledge, skills, and attitudes" (Deadorff, 2006, p. 247) teaching cultural aspects of a

language is of great importance. Ho (2009) asserted that it is impossible to separate language and culture from each other, and that it is crucial to incorporate ICC in the language teaching process. Lack of cultural knowledge can cause on learners serious problems such as disharmony, misunderstandings and even conflicts in communication in intercultural situations. This confirms that the ability to communicate successfully with people from other cultures depends not only on language skills but also on comprehension of cultural habits and expectations. Consequently, developing interculturally competent individuals, who will be cognitively aware of other cultures; who will have the ability to demonstrate appropriate behavior in interaction with people from other cultures, and who will show international attitude across cultures, becomes one of the ultimate goals of teaching a foreign language today. Thus, language teachers should plan pedagogical tasks and activities that can promote L2 learners' ICC and prepare them to meet the requirement of effectively acting in a global village (Ortiz & Moore, 2000).

Among various models of IC, which have been proposed by scholars with the purpose of describing the importance of intercultural competence from different perspectives (Arasaratnam & Doerfel 2005; Deadorff, 2006; Fantini, 2000; Chen and Starosta, 1999; Spitzberg, 1990), Byram's (1997) model of ICC was developed from a foreign language education perspective, in order to describe a person's competences required in intercultural settings. Bayram's ICC model is considered to be one of the most comprehensive frameworks to develop and evaluate learners ICC in different contexts. Byram's (1997) intercultural competence, which relates to linguistic, sociolinguistic and discourse competences, consists of five main components, which are attitudes, knowledge, skills of interpreting and relating, skills of discovery and interaction, and critical cultural awareness.

Byram (1997) describes the components of ICC by giving a detailed description of each. According to Byram (1997) the attitudes required for successful intercultural communication need to include curiosity, openness, and readiness to suspend disbeliefs and judgments about other cultures and about one's own.

Byram (1997) defines two kinds of knowledge: (1) knowledge of social groups and their practices in one's own or in one's interlocutor's country, and (2) knowledge of the process of societal and individual interaction. These kinds of knowledge are partly acquired through socialization and institutionalized learning.

Skills form the other component, which is the ability to apply knowledge and use it in different situations. The two distinct categories established are (1) skills of interpreting and relating, and (2) skills of discovery and interaction. The skills of interpreting and relating are used when individuals, based on their previous knowledge, are required to analyze, interpret and relate to a manifestation of a different culture; the skills of discovery and interaction refer to the ability to recognize important cultural phenomena, analyze their meanings and find out how they interact with other phenomena, consequently, to acquire new knowledge (Byram, 1997). In other words, the required skills include the ability of making use of existing knowledge together with the ability to recognize and acquire new knowledge in the course of the interaction.

A fourth component critical cultural awareness enables individuals to critically evaluate perspectives, practices and products of their own, and their interlocutors' cultures.

ICC assessment instruments

According to Judit (2013) empirical research studies which were carried out to find the ways for assessing ICC can be categorized as studies dealing with (1) international students' development of ICC in a foreign country, and (2) EFL students' development of ICC in a classroom environment. Also the assessment tools developed with the purpose of assessing ICC can be distinguished according to whether individuals report their experiences and perceived ICC or observers assess participants' ICC. Since our purpose is assessing students ICC in a classroom environment, with the help of a self report instrument, in this section an overview of empirical researches based on developing ICC instruments are discussed.

Assessment tools relying on individuals' self-report are indirect tools, as they do not survey actual intercultural behavior, but perceived or imagined behavior (Judit, 2013). There are many available self-report instruments designed with the purpose of assessing ICC. The example of such assessment tools of 1990's are the Intercultural Sensitivity Inventory (Bhawuk & Brislin, 1992), the Cross-Cultural Adaptability Inventory (Kelley & Meyer, 1995) and the Intercultural Development Inventory (IDI) developed by Bennett and Hammer, (1998). Intercultural Sensitivity Inventory (Bhawuk & Brislin, 1992) was designed to assess how an individual adapts to differences between living in an individualistic culture like United States and in a collectivistic culture like Japan; Cross-Cultural Adaptability Inventory (Kelley & Meyer, 1995) has also been developed to assess individual's level of adaptability to the cultures different from his/her own based on the dimensions such as emotional reliance, flexibility and openness, perceptual acuity, and personal autonomy.

The Intercultural Development Inventory (IDI) developed by Bennett and Hammer (1998) was constructed to measure the development of a person's attitude toward another culture on the basis of six stages: three ethno-centric stages (denial, defense, and minimization) and three ethno-relative stages (acceptance, adaptation, and integration). The ethno-centric orientations are applied when a person's culture is experienced as central to reality. The ethno-relative orientations are applied when a person's culture is experienced in the context of other cultures (Matveev & Merz, 2014). Generally, it aimed to measure the level of worldview orientation differences and intercultural sensitivity of the pre-service educators prior to the cultural immersion experience, before and after the process.

One more instrument which is used by many scholars for measuring ICC, is the Intercultural Sensitivity Scale developed by Chen and Starosta (2000) to measure students' ICC levels. The authors argue that intercultural sensitivity is the affective dimension of intercultural communication competence. Therefore, the scale can help individuals distinguish how their culturally different counterparts vary in behaviors, perceptions, and feelings so that they may be conscious and respectful within their interaction (Chen & Starosta, 2000). However, Arasaratnam (2009) insists that there

needs to be more research on establishing the extent to which intercultural sensitivity is a predictor of ICC. Although intercultural sensitivity may be a predictor of ICC, it is conceptually different from ICC (Arasaratnam, 2006). Derive from this, Arasaratnam (2006) developed a new instrument for measuring ICC on the basis of cognitive, affective, and behavioral components of ICC which composed of empathy, motivation, attitude toward other cultures, and interaction involvement (experience and listening) as important elements of intercultural communication competence. It also includes 'the measure of experience' part, which was based on participants studied abroad or lived abroad experiences, whether they had formal training in intercultural communication, and intercultural friendships. Another instrument Intercultural abilities scale was developed by Fantini (2005) and included four dimensions as attitudes, skills, knowledge and awareness. It was developed in order to learn explore the outcomes of intercultural service experiences – the level of intercultural competence developed by volunteers, effects on their lifestyle choices, and their impact on communities and other individuals after returning home.

In spite of the sufficient number of the instruments measuring ICC, the reason for developing a new instrument was that the above mentioned available instruments were not appropriate for the purpose of the present study. They mostly are used in order to assess learners' development of ICC during international experiences in the context of other cultures or the effect of studying abroad experiences after returning home. Thus, the main aim of this study was to develop a self-reported instrument which will measure pre-service teachers' perceived ICC levels in the classroom environment, because for the majority of participants study-abroad programs and exchange trips for international experiences are not accessible.

METHOD

This study is a part of the thesis work titled *"Predictors of Intercultural Communicative Competence in a Turkish and Kazakh Settings: Academic Self Concept and L2 Motivational Self System"*. Its purpose includes finding out differences between Turkish and Kazakhstani ELT pre-service teachers' in terms of their levels of ICC, academic self-concept and L2 motivational self system; to examine whether there is a relationship among intercultural communicative competence, academic self-concept and L2 motivational self system of ELT pre-service teachers; and which of these two affective factors can best predict the participants' levels of ICC in both two settings.

The reason why these two groups were selected was the idea that Kazakhstan and Turkey are sister nations with deep cultural and historical ties. However, Kazakhstan is a post soviet country, and country is ethnically very diverse, has a large Russian/Slavic population. Although this two countries share bonds of kinship, cultural heritage, religious and ethnic ties, Kazakh people are still influenced by Russian culture and Turkey, in its turn, is more influenced by the Western culture. Keeping in mind these similarities and differences, this study is aimed to develop and validate a scale which will be used in both Turkey and Kazakhstan in order to investigate and compare the ELT pre-service teachers ICC levels.

Research design

In the current study a survey model was used in order to develop a valid and reliable ICC scale. According to Hinkin (1995) scale development should include five basic steps which are creating an item pool, getting experts opinions, pilot study, validity and reliability analysis. All these steps required were followed to develop a valid and reliable scale in this study.

Participants

The investigation was done in two separate sections: Exploratory factor analysis and Confirmatory factor analysis. The participants of the study consisted of a total of 624 Kazakh and Turkish sophomores, junior, and senior undergraduate university students. They were all the students of English Language Teaching (ELT) departments from state universities of two countries. Since the instrument was developed in order to be applied in two settings with the purpose of comparing the two countries participants ICC, 307 English pre-service teachers from Hacettepe, Duzce and 19 Mayıs Universities in Turkey were involved in the instrument development analysis (EFA), and 314 English pre-service teachers from Akhmet Yassawi and Auezov South Kazakhstan State Universities were used for Confirmatory Factor Analysis (CFA) to test the structural validity for the developed instrument.

The participants in this study were selected through convenience sampling. This occurs when "the subjects are chose because of the close proximity to a researcher, that is, the ones that are easier for the researcher to access" (Etikan, Musa, & Alkassim, 2016, p.1)

Instrument development

Byram's conceptualization of ICC comprising knowledge, skills, attitudes and awareness was used as a theoretical basis in preparing an item pool for ICC scale. The items were constituted using five-point Likert-type scales (1 = totally disagree, 2 = disagree, 3 = partially agree, 4 = agree and 5 = totally agree).

As a main step, content validity analysis was conducted in order to evaluate the scale items on the relevancy, clarity and conciseness of the items. After content validity scores were calculated and necessary corrections were made, a pilot implementation was carried out on 314 voluntary student-teachers studying at Hacettepe, Duzce and 19 Mayıs Universities. According to the criterions proposed by Tabachnick and Fidell (2007) a population size for factor analysis is considered "fair" for 200 people; "good" for 300 people; "very good" for 500 people; and "excellent" for 1,000 people. Consequently, the sampling employed for this study was considered good.

Exploratory factor analysis was conducted to test the structure validity and to make clear the dimensions of a new ICC scale. Kaiser-Meyer-Olkin (KMO) coefficient and Barlett Sphericity test were calculated in order to check the suitability of the data for factor analysis. Alpha internal consistency coefficient and item total correlations were calculated to check the reliability of each sub-dimension. Confirmatory factor analysis was used to test whether a new developed ICC scale fits to Kazakh context. SPSS 22.00 software program was used for EFA and Lisrel 8.70 program was used for CFA.

FINDINGS

Content validity

Content validity refers to the extent to which the items on a measure assess the same content or how well the content material is sampled on the measure. Content validity includes face validity and logical validity. Face validity indicates whether the measure is valid "on its face" while logical validity indicates a more rigorous process, such as using a panel of experts to evaluate the content validity of a measure (Rubio, et. al., 2003).

In order to evaluate the item pool which consists of 74 items on the relevancy, clarity and conciseness, 12 purposely chosen experts were asked to review the draft. This procedure is entailed confirmation by a specific number of experts, indicating that instrument items and the entire instrument have content validity. The experts consisted of totally 12 academicians from Turkey and Kazakhstan: 3 professors, 5 Associate Professors, 3 Ph Doctors and 1 academician with master's degree.

The consensus among experts on the necessity to include a specific component was quantified by determining the content validity ratio for each item. With this purpose, the experts were requested to specify whether an item is necessary for operating a construct in a set of items or not. They were asked to rate each instrument items in terms of clarity and its relevancy to the construct underlying study as per the theoretical definitions of the construct itself and its dimensions from 1 to 3 with a three-degree range of (1) relevant; 2) relevant but needs correction; 3) not relevant. Content Validity Ratio (CVR) varies between +1 and -1. The higher score indicates further agreement of members of the panel on the necessity of an item in an instrument. The formula of content validity ratio is CVR=Ne / (N / 2) - 1,   in which the Ne is the number of panelists indicating "Relevant" and N is the total number of panelists. The numeric value of content validity ratio is determined by Lawshe Table. In our study since the number of panelists 12 members, if CVR is bigger than 0.56, the item in the instrument with an acceptable level of significance will be accepted.

As a next step, the Content Validity Index (CVI) was computed to estimate the validity of the whole instrument. It presents the commonality of judgments regarding the validity or applicability of the final procedure, model, test, or format being researched. The overall content validity will be higher if the value of the CVI is closer to 0.99 and vice versa (Lawshe, 1975). Content validity Index was calculated by *CVI= Σ¹SVR/ RETAINED NUMBER.* The CVI for the ICC scale was found to be 0.79 in the current study, which showed the acceptable validity of a whole instrument. According to the CVI results, none of the items were rejected but corrections were made based on the experts' recommendations.

Exploratory Factor Analysis

Exploratory Factor analysis is a statistical method commonly used during instrument development to cluster items into common factors, interpret each factor according to the items having a high loading on it, and summarize the items into a small number of factors (Bryman & Cramer 1999). Loading refers to the measure of association between an item and a factor (Bryman & Cramer 2005). A factor is a list of items that belong

together. Related items define the part of the construct that can be grouped together. Unrelated items are those that do not belong together, do not define the construct and should be deleted (Munro 2005).

The suitability of the data for factor analysis was calculated by applying Kaiser-Meyer Olkin (KMO) sampling adequacy and the Bartlett's test of sphericity. The KMO statistic varies between 0 and 1. A value of 0 indicates that the sum of partial correlations is large in the sum of correlations, which indicates diffusion in the pattern of correlation, and that factor analysis is inappropriate. A value close to "one" indicates factor analysis will yield distinct and reliable factors (Field 2009).

Table 1
KMO and Barlett Tests results

Kaiser-Meyer-Olkin sampling adequacy test		,941
Bartlett's Test of Sphericity	Approx Chi-Square	14984,992
	df	2850
	sig	,000*

p<.05*

According to the results, The Kaiser-Meyer-Oklin value of the scale was calculated as .939, indicating that the sampling is highly adequate. Bartlett's test results showed X^2=14570,091; sd=2850 (p<.000). It is acceptable if the KMO measurement test result is .60 and over, and the result of the Barlett Sphericity test is statistically significant with a minimum acceptable coefficient which is .60 (Tabachnick and Fidell, 1996). So these two values pointed out the data set is eligible for factor analysis.

In order to reveal the factor design of the instrument, principal component analysis and varimax rotated component matrix were chosen as the factor analysis. The lower cut-off point of the factor loads was taken as 0.45, and those with loads lower than .45 were removed from the instrument. Varimax rotation showed that the scale had fourteen factors higher that 1.

Concerning to the results obtained from factor analysis, the variance ratio of four components altogether explained 46.48% of the variance. From the fifth factor, the eigenvalues of factors were found to be close to each other. These results therefore showed that the scale had a structure with four factors. Keeping the number of factors high increases the explained variance, but this time it is likely to be difficult to name the factors and mak them meaningful (Büyüköztürk, 2010). The greater the variance ratios obtained at the end of the analysis, the stronger the factor structure. This level is considered to be between 40% and 60% in social sciences (Tavşancıl, 2002).

Although the results of principal component analysis determined 14 factors, it showed that the eigenvalues of factors from 5 to 15 were very close to each other and there were not enough items in these factors. After omitting 14 items with loadings lower than .45, the data was reanalyzed with the rest of 62 items in order to create a four dimensional factor structure. According to the results of reanalyzed data, PCA based on four factors explained 47,780% of the variance. After removing 12 items whose factor loadings were less than .45 and overlapping with more than one factor, PCA was conducted again.

Finally, the results revealed that the variance ratio explained 50.48% of the variance (the first factor explained 34.430 of total variance, the second one explained 7.446% of it, the third factor 5.067%, and the forth factor explained 3.742% of total variance). According to Kline (1994), 40% of the total variation explained by factors can be taken as the minimum acceptable explained variance level.

Ultimately, a valid and reliable measurement tool with 52 items which grouped under four factors was developed. The first factor consisted of 21 items associated with intercultural skills, the second factor consisted of 13 items associated to intercultural attitudes, third factor associated to intercultural awareness (12 items), and fourth factor associated to intercultural knowledge (6 items). The Table 2 represents the results of PCA in terms of means, standard deviations, factor loadings and item total correlation values of ICC scale.

Table 2
Validity analysis results of ICC scale

Items	Mean	Standard Deviation	Component Factor Loadings	Varimax Factor Loadings	Item Total Correlations
Factor 1					
30. I am able to express my thoughts and ideas clearly when interacting with people from different cultures.	3,74	,878	,541	,536	,648
32. I am able to use appropriate body language when interacting with people from different cultures.	4,03	,829	,557	,563	,644
33. I am able to interact and communicate effectively with people from different cultures.	3,92	,832	,670	,663	,722
34. I am able to communicate appropriately by taking into consideration norms and beliefs of people from different cultures.	4,02	,812	,590	,589	,639
35. I am able to help my friends and classmates to solve cross-cultural misunderstandings when they arose in any situations.	3,86	,888	,600	,600	,650
36. I am able to initiate a conversation when I meet people from different cultures.	3,75	,962	,600	,649	,692
37. I am able to keep going a conversation during the interaction with people from different cultures.	3,87	,851	,673	,675	,774
38. I am able to interact/ communicate appropriately in the markets, shops and other public places with people from different cultures.	4,00	,828	,511	,529	,653
43. I can cooperate easily with people from different cultures on shared activities and ventures.	3,85	,811	,650	,649	,720
47. I can follow all grammar rules when interacting with people from other cultures.	2,93	1,067	,514	,519	,324
50. I am able to manage breakdowns in communication with people from different cultures.	3,52	,857	,691	,690	,689
51. I can deal with problems by my own in foreign counties.	3,75	,972	,549	,546	,629
56. I am able to make an intercultural friendship.	4,01	,814	,484	,479	,684
58. I am able to solve problems stemming from cultural differences.	3,61	,876	,642	,627	,706
59. I am able to identify differences and similarities across my own and other cultures.	4,04	,744	,476	,467	,560
60. I am able to deal with culturally distinct persons.	3,71	,889	,656	,652	,636
64. I am able to initiate and terminate conversation appropriately with people from other cultures.	3,74	,830	,669	,662	,736
65. I am able to maintain the communication with people from other cultures.	3,89	,823	,638	,648	,748
66. I am able to express myself clearly when the situation requires it.	3,95	,821	,623	,636	,708
73. I am confident when interacting with people from different cultures.	3,65	,937	,588	,584	,641

74. I can use appropriate verbal behavior (e.g. accent, tone) when communicating with people from other cultures.	3,66	,946	,553	,550	,677
Factor 2					
2. I am willing to learn about other cultures' traditions and norms.	4,50	,635	,600	,612	,679
3. I am willing to communicate with people from other cultures (who have different perceptions and orientations from mine).	4,47	,692	,705	,703	,676
4. I am interested in meeting people from different cultures and countries.	4,46	,715	,749	,754	,691
5. Interacting with people from different cultures makes me happy.	4,49	,655	,738	,740	,691
7. I always try to come into contact with people from other cultures when it is appropriate.	3,84	,,905	,548	,555	,515
11. I would like to have a lot of friends from different cultures.	4,14	,975	,502	,501	,599
16. I get a lot of pleasure from taking part in different intercultural activities such as music festivals, fairs, concerts etc.	4,06	1,00	,612	,604	,694
17. I like visiting music festivals and concerts of different cultures.	3,96	1,00	,586	,583	,638
18. I like visiting fairs of different cultures.	3,91	,978	,612	,609	,633
19. I am eager to visit theatrical plays of different cultures.	3,90	1,02	,566	,566	,591
22. I am willing to take part in different intercultural educational and scientific projects.	4,04	,960	,568	,563	,600
23. I would like to join in different intercultural courses and programs abroad.	4,21	,872	,613	,607	,576
28. I am eager to make friends from different cultures and countries.	4,13	,879	,568	,566	,595
Factor 3					
8. I feel nervous when interacting with my people from other cultures at school.	3,47	1,057	,707	,709	,684
9. I find it difficult to tell the direction to foreigners.	3,65	1,022	,602	,608	,571
10. I often get confused when it is my turn to express myself in front of people from other cultures.	3,29	1,031	,608	,619	,543
12. I find it difficult to make friends from other cultures.	3,56	1,022	,499	,505	,341
26. I feel myself uncomfortable while interacting with people from other cultures.	3,63	,986	,692	,673	,652
29. I find it difficult to get into contact with people from different cultures.	3,62	,975	,696	,695	,654
41. It is difficult for me to deal with behaviors of people from other cultures.	3,58	1,00	,587	,594	,603
42. I am able to express my thoughts in written form by using a second language.	3,64	,969	,694	,703	,669
54. I am able to express my thoughts in written form by using a second language.	4,27	1,091	,473	,462	,551
69. I feel anxious when communicating with people from different cultures.	3,29	1,157	,627	,636	,571
75. I do not feel confident enough to make friends from other cultures.	3,76	1,052	,618	,634	,627
76. My language competence is not enough for interacting with people from other cultures.	3,83	1,04	,571	,583	,475
Factor 4					
14. I get pleasure from listening to the music of different cultures.	4,36	,855	,571	,597	,623
15. I like watching films of different cultures.	4,44	,774	,545	,565	,524
39. I am able to read, understand and interpret books, magazines, articles etc., of different cultures.	4,25	,782	,637	,658	,552
40. I am able to understand the advertising boards and road signs when visiting foreign countries.	4,04	,922	,538	,560	,522
45. I know about the importance of other cultures' values and beliefs in communicating with people from different cultures.	4,27	,760	,529	,544	,601
61. I am interested in different topics such as films, music, art, etc. of different cultures.	4,34	,832	,601	,620	,574

It was found that item-total correlations of all items ranged from .341 to .774. While item-total correlations vary from .476 to .691 for Factor I, for Factor II correlations vary from .502 to .749, for Factor III it ranged from .473 to .707, and for the last factor it ranged from .523 to .637. Considering that items with item-total correlations above .30

discriminated well (Büyüköztürk, 2005), it can be stated that the items measure the same behavior at a fair or high level.

Reliability analysis

The Cronbach alpha coefficient is used in order to measure the internal consistency of the scale. To determine the reliability of the instrument, Cronbach alpha values for each factor was computed. the reliability coefficient ranges from 0 and 1, the closer the value of alpha to 1.0, the greater the internal consistency of the instrument (Gliem & Gliem,2003). However, it is suggested that an alpha coefficient of 0.70 is acceptable for a new instrument (DeVellis 2003; DeVon et al. 2007).

Table 3
Reliability Cronbach's alpha coefficients of four factors

Components	Skills	Attitude	Awareness	Knowledge	Total:
α	,946	,906	,880	,806	,958

Cronbach's alpha coefficients of the factors were found as .946; .906; .880 and .806, respectively, which shows a high reliability level of each dimension. Totally, the Cronbach alpha reliability coefficient for the whole instrument was calculated as .958. So the results from the reliability analysis indicated a high correlation between the items and showed that the instrument is consistently reliable.

Table 4
Correlations results between factors and total instrument items

Dimensions	N	X	Std	p	Skills	Attitude	Awareness	Know ledge	Total
Skills	314	3,79	,604	,000*	-	,672**	,590**	,563**	,928**
Attitude	314	4,16	,602	,000*	,672**	-	,439**	,586**	,825**
Awareness	314	3,63	,680	,000*	,590**	,439**	-	,278**	,752**
Knowledge	314	4,28	,586	,000*	,563**	,586**	,278**	-	,656**
Total	314	3,90	,513	,000*	,928**	,825**	,752**	,656**	-

P<.05*

When the correlation coefficients between factors were analyzed, Pearson's coefficient of correlation indicated high levels of positive correlations of all dimensions with overall ICC scale. It can be seen from the table that there is a significant inter-correlation between factors and a total scale which ranges from .656 to .928.

Confirmatory Factor Analysis

The final version of ICC scale with total of 52 items was conducted on 307 pre-service EFL pre-service teachers' of two universities in Kazakhstan with the aim of confirmation the four factorial structure of the instrument.

In order to test the structure validity of the instrument and accuracy of sub-dimensions obtained by the Exploratory Factor Analysis, Confirmatory Factor Analysis (CFA) was applied using LISREL 8.7. CFA is a version of factor analysis in which specific hypotheses about structure and relations between the latent variables that underlie the data are tested (Field, 2009). CFA is not concerned with discovering a factor structure, but with confirming the existence of a specific factor structure. It is used in later phases

of scale development or construct validation after the underlying structure has been tentatively established by prior empirical analyses using EFA, as well as on theoretical grounds (Brown, 2006).

The technique of CFA analyzes a priori measurement model in which both the number of factors and their correspondence with the indicators are explicitly specified (Kline, 2011). Table 5 shows error and fit index in items.

Table 5
Error and Fit Index for ICC scale

Fit Index	Acceptable Fit	Suggested New Model
$\chi2/df$	$.00 < \chi2/sd < 3$	2353.90/1263=1.8
RMSEA	$.05 \leq$ RMSEA $\leq .10$	0.58
RMR	$.00 \leq$ RMR $\leq .10$	0.056
SRMR	$.00 \leq$ SRMR $\leq .10$	0.065
NFI	$.90 \leq$ NFI $\leq .95$	0.93
NNFI	$.95 \leq$ NNFI $\leq .97$	0.96
CFI	$.90 \leq$ CFI $\leq .95$	0.96
GFI	$.90 \leq$ GFI $\leq .95$	073
AGFI	$.80 \leq$ AGFI $\leq .90$	071
PGFI	$.00 \leq$ PGFI $\leq .95$	067

Factor structure of the new instrument, which has four subdimensions with 52 items, was tested with the CFA. First of all, fit indices were examined to evaluate the overall fit. Chi Square ($\chi2$), RMSEA, NFI, NNFI, CFI, GFI, and AGFI are the most commonly used statistical analysis in model data fit structure.

The chi-square goodness-of-fit ($\chi2$) statistic was statistically significant but the chi-square statistic is sensitive to the sample size, so it is rarely used as a sole index of the model fit. An adjunct discrepancy based fit index is the ratio of chi-square to degrees of freedom ($\chi2/df$). If that ratio is in the range of 2 to 3, it is indicative of an acceptable fit between the hypothesized model and the sample data (Tabachnick & Fidell, 2007). Results indicated $\chi2 = 2353,90$ df = 1263, $\chi2/df = 1.8 < 2$, p=0.000, Goodness of Fit Index (GFI =.73), The Adjusted Goodness of Fit Index (AGFI = .71), The Normed Fit Index (NFI=.93), The Non-Normed Fit Index (NNFI = .96), and the comparative Fit Index (CFI=.96). All these incremental indices are scaled from 0 (no fit) to 1 (perfect fit). Researchers (Tabachnick & Fidell, 2007; Kline, 2011) advise that values close to 0.95 are indicative of a good fit. According to the results of the study, almost all indices were close to 0.95. The root mean square error of approximation (RMSEA) is a population discrepancy function that compensates for the effects of the model's complexity. The closer the RMSEA coefficient is to 0, the better the fit of the model. The RMSEA value of .05 or less indicates a close fit of the model in relation to the degrees of freedom, whereas a value of .08 or less indicates a reasonable error of approximation (Tabachnick & Fidell, 2007). In this study RMSEA was computed as 0.058. According to all these results of CFA the model consisted of four factors provided a good model fit for using in Kazakh context to examine EFL pre service teachers ICC. Figure 1 below reveals the factor distribution and the interaction among the subscales.
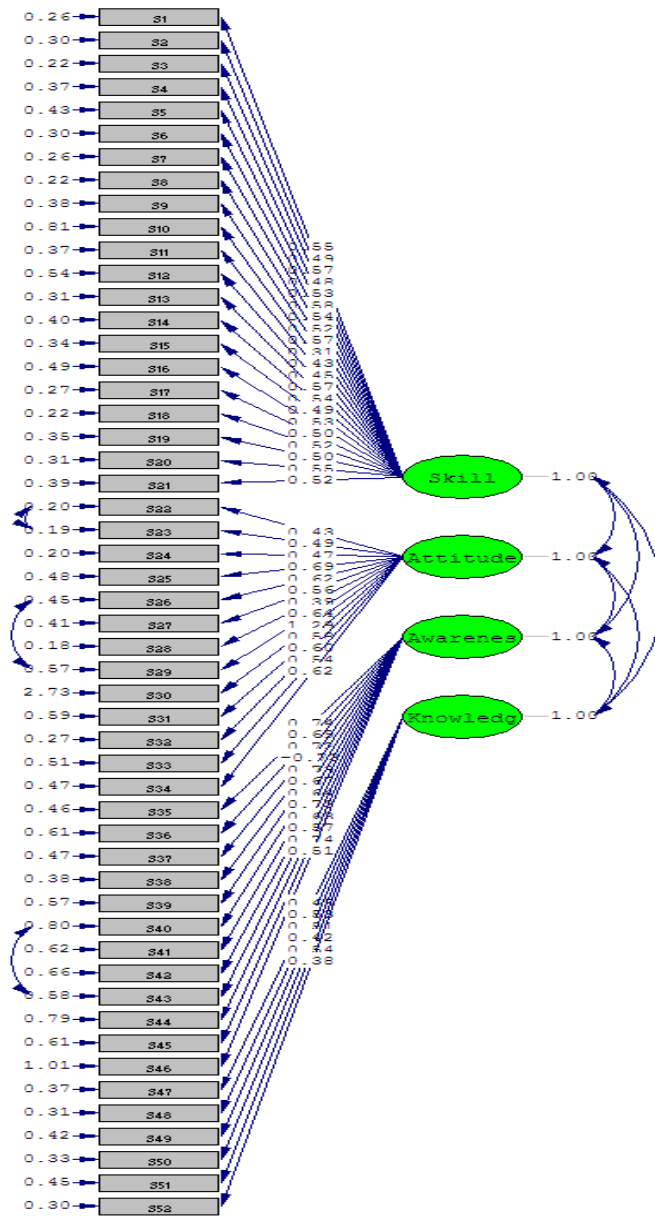
Figure 1
Factor distribution and the interaction among the subscales

Reliability analysis

The internal consistency of the confirmed ICC scale was computed for all four dimensions'. The results are indicated in the table below.

Table 6
Cronbach's alpha results of four factor instrument after CFA

Components	Skills	Attitude	Awareness	Knowledge	Total
α	,935	,860	,807	,764	,937

The Cronbach's Alpha scores for each sub-dimension were as follows: Skills = .935, Attitude = .860, Awareness =.807, Knowledge = .764, and finally a Reliability Coefficient of a whole instrument indicated .937, which showed an acceptable index. Cronbach α value above 0.60' shows that the instrument is reliable (Kalaycı, 2010).

Table 7
Correlations results between factors and total instrument items after CFA

Factors	N	X	Std	p	Skills	Attitude	Aware ness	Knowle dge	Total
Skills	258	3,88	,519	,000	-	,609**	,416**	,605**	,897**
Attitude	258	4,13	,603	,000	,609**	-	,282**	,541**	,794**
Awareness	258	3,58	,590	,000	,416**	,282**	-	,367**	,654**
Knowledge	258	4,25	,519	,000	,605**	,541**	,367**	-	,723**
Total	258	3,92	,440	,000	,897**	,794**	,654**	,723**	-

P<.05*

In order to confirm the applicability of the instrument to the Kazakh context, factor correlations and item total correlations were calculated. The results revealed that correlation between factors and a total scale ranges from .654 to .897, which demonstrated a significant positive correlation between factor and overall scale.

DISCUSSION

The aim of this study was to develop a self-assessment tool that measures ELT pre-service teachers' intercultural communicative competence levels. After reviewing a broad existing relevant literature on the ICC, Byrams' model of ICC was taken as a basis for preparing an item pool. To establish the validity, twelve experts judged the instrument on the relevancy, clarity and conciseness of the items. Content validity index of each item was calculated based on the experts' opinions. According to the CVI results none of the items were rejected but made corrections based on the experts recommendations. Further, a draft form consisting of 76 items conducted on pre-service teachers of English language departments in Turkey and Kazakhstan. EFA and CFA were used to analyze the structural validity of the scale. As a result of EFA 24 items were omitted from the scale and a structure consisting 52 items and four factors were obtained. The factor loading values of the items varied between .324 and .774 and explained 50.48% of the total variance.

As a result of EFA the final ICC Scale was grouped under four headings:

- "ICC skills" subscale included 21 items which were based on Byrams' conceptualization assessing participants abilities to use an appropriate combination of attitudes, skills and knowledge to interact with interlocutors from a different country and culture, taking into consideration the degree of one's existing familiarity with the country and culture and the extent of difference between one's own and the other (Byram, 1997).

- "ICC Attitude" subscale included 14 items concerning with the participants attitudes towards people from different countries, different cultures, traditions and norms.

- "ICC Awareness" subscale is consisted of 12 items to assess the participants' critical evaluation of themselves, which requires them to be aware of their own attitudes, beliefs, values and perceptions.

- "ICC Knowledge" subscale includes 6 items which assess the participants' knowledge about other countries, and their beliefs, meanings and behaviors' (Byram, 1997).

Further, the factors obtained from EFA were tested with CFA in order to confirm the factor structure of ICC scale. The results of CFA demonstrated that the instrument consisted of four factors provide a good model fit. Moreover, reliability coefficients were calculated for both two contexts which exhibits high degree of internal consistency (Turkey α=.958; Kazakhstan α=.937). Finally, it can be said that the scale passed all criteria such as content and face validity, structural validity and reliability analyses.

CONCLUSION

Researchers have used highly valid and reliable scale development procedures but some limitations of the study may provide opportunities for further studies in the future. The main limitation is the sample used. Since students of ELT program were used in this study, the students of different programs may provide different results compared to the used sample. Also, since there are no items based only to ELT students, there seems to be no reason why it should not be suitable for all EFL learners. Therefore, a more random and bigger sample size from different locations with participants of varying age, gender, and education level can be used to strengthen the outcome of the study.

REFERENCES

Arasaratnam, L. A. (2006). Further testing of a new model of intercultural communication competence. *Communication Research Reports*, *23*(2), 93-99.

Arasaratnam, L. A. (2009). The development of a new instrument of intercultural communication competence. *Journal of Intercultural Communication*, (20).

Arasaratnam, L. A., & Doerfel, M. L. (2005). Intercultural communication competence: Identifying key components from multicultural perspectives. *International Journal of Intercultural Relations*, *29*(2), 137-163.

Bhawuk, D. P., & Brislin, R. (1992). The measurement of intercultural sensitivity using the concepts of individualism and collectivism. *International journal of intercultural relations*, *16*(4), 413-436.

Bennett, M. J., & Hammer, M. R. (1998). The intercultural development inventory (IDI) manual. *Portland, OR: The Intercultural Communication Institute*.

Brown, T.A. (2006). *Confirmatory factor analysis for applied research.* The Guildford Press: London.

Bryman, A., & Cramer, D. (1999). Quantitative data analysis with SPSS release 8 for Windows. *A guide for social scientists. London and New York: Taylor & Francis Group.*

Bryman, A., & Cramer, D. (2005). Quantitative analysis with SPSS 12 and 13: A guide for social scientist.

Büyüköztürk, Ş. (2005). *Sosyal Bilimler İзin Veri Analizi El Kitabı.* Ankara: Pegem A Yayıncılık..

Büyüköztürk, Ş. (2010). *Data analysis for social science hand book (11th edition)*, Ankara: PegemA Publishing.

Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Multilingual Matters.

Chen, G. M., & Starosta, W. J. (2000). The development and validation of the intercultural sensitivity scale.

Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a

student outcome of internationalization. Journal of Studies in International Education 10(3): 241-266.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage.

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D.J., et al.

(2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship, 39*, 155-164.

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1-4.

Fantini, A. E. (2005). Assessing intercultural competence: a research project of the federation EIL. *SIT Occasional Papers Series*.

Field, A. (2009). *Discovering statistics using SPSS*. Sage publications.

Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education.

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21*(5), 967-988.

Ho, S. T. K. (2009). Addressing culture in EFL classrooms: The challenge of shifting from a traditional to an intercultural stance. *Electronic journal of foreign language teaching*, *6*(1), 63-76.

Judit, D. (2013). A Mixed-Method Study on English Majors' Intercultural Communicative Competence.

Kalaycı, Ş. (2010) SPSS UygulamalıÇokDeğişkenliİstatistikTeknikleri, ÖzBaranOfset, 5.Baskı: Ankara.

Kelley, C., & Meyers, J. (1995). *CCAI: Cross-Cultural Adaptability Inventory: Manual*. National Computer Systems, Incorporated.

Kline, R.B. (2011). *Principles and Practice of Structural Equation Modeling, Third Edition.*, New York:The Guilford Press.

Kline, P. (1994). An easy guide to factor analysis. New York, NY: Routledge.

Mažeikienė, N., & Virgailaitė-Mečkauskaitė, E. (2007). The experience of measurement and assessment of intercultural competence in education. *Social Sciences,* 58, 70-82.

Matveev, A. V., & Merz, M. Y. (2014). Intercultural Competence Assessment: What Are Its Key Dimensions Across Assessment Tools?. In *Toward Sustainable Development Through Nurturing Diversity: Selected Papers from the Twenty-First Congress of the International Associate for Cross-Cultural Psychology* (pp. 141-153).

Munro, B. H. (2005). *Statistical methods for health care research*(Vol. 1). Lippincott Williams & Wilkins.

Rubio, D. M., Berg-Weger, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social work research*, *27*(2), 94-104.

Spitzberg, B. H. (2000). A model of intercultural communication competence. *Intercultural communication: A reader*, *9*, 375-387.

Tabachnick BG, Fidell LS (2007). Using Multivariate Statistics. Pearson Education. Boston, MA.

Tavşancil, E. (2002). Tutumların ölçülmesi ve SPSS ile veri analizi. *Nobel Yayıncılık, Ankara*.

Yurdugül, H. (2005). *Content validity index usage in scale development studies for content validity,* XIV. National Congress of Educational Sciences, p.1-6, September 28-30, Denizli, Turkey.