NOVITAS-ROYAL
RESEARCH ON YOUTH AND LANGUAGE

2022, 16(1), 74–96
Novitas-ROYAL (Research on Youth and Language)
ISSN: 1307–4733

Research Article

Adaptation and Validation of the Intercultural Sensitivity Scale to Turkish EFL Contexts and for Young Learners[1]

Özge Pelin NAZLI[2] & Asuman AŞIK[3]

[2]M.A., Gazi University, Ankara, TURKEY
ozgepelinnazli@gmail.com
ORCID : 0000-0002-9430-6857
[3]Assoc. Prof. Dr., Gazi University, Ankara, TURKEY
asuman.asik@gazi.edu.tr
ORCID: 0000-0003-3293-1283

Article information

Submission	30/12/2021	Revision received	29/03/2022
Acceptance	15/03/2022	Publication date	24/04/2022

Keywords: Intercultural sensitivity, young learners, scale adaptation and validation

Abstract: Interculturality and intercultural communicative competence were accepted as significant concepts in second/foreign language learning. In particular, intercultural sensitivity (IS) has been regarded as a precondition for intercultural communicative competence and interculturality. Due to the increased interest in intercultural sensitivity and globalization through technology, more research was needed to explore the concept in different educational contexts with sound, reliable and valid instruments for different participant profiles. In this regard, this study merged two studies to provide two reliable and valid instruments for different research purposes. The objective of Study 1 was to translate and validate the IS scale in the Turkish EFL context, while the objective of Study 2 was to adapt and validate the Turkish version of the IS scale to be used for young English language learners in the Turkish EFL context. Study 1 included 122 language learners, while Study 2 included 95 young learners for the first round and 115 young learners for the second round. The findings revealed that both the Turkish version and the simplified version of the IS scale retained highly significant levels of reliability and validity. Confirmatory factor analyses indicated that the Turkish version represented a better model fit in the analysis and the simplified version represented a better model fit compared to the English version in the analysis. Hence, the study is significant with its contribution to the field through both the Turkish version of IS and IS scale for young learners, which can be used in cross-cultural research contexts.

Anahtar Sözcükler: Kültürlearası duyarlılık, çocuklar, ölçek uyarlama ve geçerliği

Kültürlerarası Duyarlılık Ölçeğinin Türkiye Yabancı Dil Öğrenme Ortamlarına ve Çocuklar için Uyarlanması ve Geçerliliğinin Sınanması

Özet: Kültürlerarasılık ve kültürlerarası iletişim yetisi, ikinci/yabancı dil öğrenmedeki önemli kavramlar olarak kabul edilmiştir. Özellikle, kültürlerarası duyarlılık, kültürlerarasılık ve kültürlerarası iletişim yetisi için gerekli bir önkoşul olarak tanınmaktadır. Kültürlerarası duyarlılığa artan ilgi ve teknolojiyle gelen globalleşme nedeniyle, etkili, güvenilir ve geçerli araçlar kullanarak farklı eğitimsel bağlamlarda bu kavramın araştırılması için daha fazla araştırmaya ihtiyaç duyulmaktadır. Bu kapsamda, bu çalışma farklı araştırma amaçları için kullanılabilmesi amacıyla iki güvenilir ve geçerli araç sunmak amacıyla iki çalışmayı birleştirmiştir. Birincisi, IS (Kültürlerarası Duyarlılık) ölçeğini Türkiye bağlamına uyarlamak ve geçerliliğini sınamaktır, ikincisi ise Türkiye dil öğrenimi bağlamındaki çocuklar için IS ölçeğin uyarlanması ve geçerliliğini sınamaktır. Birinci çalışmaya 122 yetişkin öğrenci dahil olurken, ikinci çalışmanın birinci kısmına 95 ikinci kısmına ise 115 çocuk dahil olmuştur. Bulgular, hem Türkçe versiyonun hem de basitleştirilmiş versiyonun oldukça yüksek güvenilirlik ve geçerliliğe sahip olduğunu göstermiştir. Doğrulayıcı faktör analizleri, Türkçe versiyonun daha iyi bir model olduğunu, ve basitleştirilmiş versiyonun da İngilizce ölçekten daha iyi modeli temsil ettiğini ispatlamıştır. Dolayısıyla, bu çalışma IS ölçeğinin hem Türkçe olarak hem de erken yaştaki öğrenciler için olan versiyonun, kültürlerarası araştırma bağlamlarında kullanılabilir olması dolayısıyla alana katkı sunan öneme sahiptir.

To Cite This Article: Nazlı, Ö. P., & Aşık, A. (2022). Adaptation and validation of the intercultural sensitivity scale to Turkish EFL contexts and for young learners. *Novitas-ROYAL (Research on Youth and Language)*, *16*(1), 74–96.

[1] *This article reports a part of a Master's thesis called "Telecollaborative Digital Storytelling: A Task for Enhancing Young EFL Learners' Intercultural Communicative Competence" of the first author*

1. Introduction

The impact of multiculturalism and the advent of computer-mediated communication in the 21st century has increased interest and the necessity of intercultural communicative competence (ICC) for foreign language learning. ICC has been acknowledged as a fundamental competence in second/foreign language learning as it involves the ability to be conscious of one's own culture and humanity, to develop a capacity to inquire about the social conditions in his and others' contexts, and to be capable of interacting with other cultures effectively (Byram, 1997). From a different perspective, Deardorff (2006) perceives ICC in a continuum and highlights that it commences with respect, openness, and curiosity and includes self-consciousness and abilities to listen, examine, criticize, assess, and associate with others. These abilities contribute to the development of effective communication skills and appropriate attitudes in intercultural circumstances externally. Moreover, it renders the establishment of empathy and flexibility within the ethnorelative mindset internally (Deardorff, 2006). In relation to ICC, the concept of intercultural sensitivity (IS) is also regarded as a critical aspect since it encompasses the affective dimension (Alaei & Nosrati, 2018; Dong, Day, & Collaço, 2008; Peng, 2006).

1.1. Intercultural Sensitivity

The increasing need for ICC and interculturality has led to the emergence of various theories and frameworks, some of which attach great importance to IS as a critical skill. IS refers to understanding, distinguishing, and esteeming the cultural variances during cross-cultural interactions (Chen & Starosta, 1996). It is conventionally regarded as acting attentive towards foreign cultures, recognizing and appreciating cultural diversities, and voluntarily altering one's manners out of appropriateness and respect (Hammer, Bennett & Wiseman, 2003).

As for a fundamental theoretical framework of IS, Bennett (1986) developed a comprehensive theory called The Developmental Model of Intercultural Sensitivity (DMIS). Bennett (1986) introduced a sequence of stages in the DMIS to enable educators to determine their learners' present sensitivity levels and adapt materials in line with their sensitivity levels to progress to the next stages in the developmental continuum. Designating stages in a hierarchical organization, the DMIS model allowed instructors to pinpoint the sensitivity levels of learners, choose suitable learning materials, and put these materials in a logical order to assist the progress towards higher sensitivity levels (Bennett, 1986)

Having cognitive constructivist orientation, the DMIS determines six stages of IS with incremental order and attaches certain cognitive, attitudinal, and behavioral characteristics to each stage in the continuum (Garrett-Rucks, 2014). All six stages in the model include a mindset with consecutively higher competence to process cultural variations. The former three stages in the model (Denial, Defense, and Minimization) are labeled as ethnocentric perspectives, whereas the latter three stages (Acceptance, Adaptation, and Integration) are defined as ethnorelativist perspectives. Ethnocentrism indicates the way of thinking in which individuals perceive their cultural principles as the finest of reality and evaluate others' cultural norms (Berry, Poortinga, Breugelmans, & Sam, 2011; Munezane, 2019). Ethnorelativism, on the other hand, represents the point of view in which individuals unconsciously welcome cultural variations and evaluate their vision of the world based on these variations (Çiftçi, Karaman, & Daloğlu, 2020). Gaining intercultural understanding adheres to the development from ethnocentric mindset to ethnorelativist mindset, which means stepping out from one's monocultural zone, perceiving events from others'

viewpoints, and cultivating empathy towards and awareness of transcultural exchanges with nonnative communities (Rantz & Horan, 2005). Bennett (2011) also pinpoints that intercultural awareness and transcultural interaction do not directly result in improvement in learners' ICC.

Scholars have discussed the relationship between IS and ICC from several perspectives. According to Bennett (2004), IS is of significant importance to raise interculturally competent individuals since the enhancement in learners' IS yields to their development of ICC, which represents one's ability to think and behave appropriately during transcultural interactions. From a critical viewpoint, Hammer et al. (2003) underlined the existence of a concept error regarding the function of ICC; thus, they distinguished between IS and intercultural competence to amend the misconception. From their point of view, IS is the recognition and experience of cultural differences, whereas intercultural competence is the skill of behaving and contemplating in an interculturally appropriate manner.

With a more affective and cognitive understanding, IS is learners' willingness to comprehend, appreciate and acknowledge the intercultural differences (Chen & Starosta, 1998). On the other hand, communicating in an additional language, constructing a social relationship with a person, and developing effective communication, although that person has a different mindset, represent intercultural communicative competence. (Douglas & Rosvold, 2018). Furthermore, IS is defined as an ability to move beyond his ancestry culture and communicate with interlocutors from different socio-cultural backgrounds (Sinicrope, Norris, & Watanabe, 2007). It indicates one's capacity to process and distinguish opposing cultural perceptions (Hammer et al., 2003).

Regarding the ICC and IS, Chen and Starosta (1996) suggested a framework by integrating three spheres under ICC which are intercultural awareness (e.g., the cognitive dimension of ICC), intercultural androidness (e.g., the behavioral dimension of ICC), and IS (e.g., the affective dimension of ICC). Among these spheres, IS is described as one's mental condition to cultivate positive feelings towards comprehending and welcoming cultural divergences to strengthen convenient and competent demeanor in them (Chen & Starosta, 1998). Namely, IS signifies the fitness to change one's attitude to respect individuals with culturally-distinct origins. The development in intercultural knowledge and skills does not automatically yield to critical cultural awareness and shift in the perspective, which represent the core elements of ICC. Individuals may have broad intercultural knowledge and high language proficiency to interact with a foreign culture and its members, yet they may still hold a defensive attitude towards them. Therefore, IS is viewed as a precondition to foster ICC (Hammer et al., 2003) and improve critical cultural awareness since higher IS refers to higher readiness to alter negativity in attitude and behavior and effectuates higher ICC (Dong et al., 2008).

Based on the abovementioned importance of the IS, several researchers endeavored to develop IS scales to determine learners' transcultural awareness and attitudes. However, Chen and Starosta (1996) criticize that the previous academic productions conducted to measure learners' ICC did not build their scales on a theoretical framework. Hence, the authors called for conceptual clearness. They proposed an IS framework that fundamentally concentrates on learners' emotions by unifying cognitive, attitudinal, and behavioral patterns in a single system. Chen and Starosta (1998) also pinpointed four individual features in their measure: self-concept, open-mindedness, non-judgemental, and social relaxation. Self-concept attributes to one's positive viewpoint in transcultural communications. Open-mindedness indicates his eagerness to communicate his views honestly and to acknowledge

others' expressions. The non-judgmental dimension signifies having no prejudices towards foreign cultures and paying attention to others in a candid fashion during cross-cultural interplays. The last dimension of this scale, social relaxation refers to one's skill to eliminate ambiguous emotions during multicultural dialogues. Drawn on these features, Chen and Starosta (2000) developed a 24-item IS scale which encompasses five factors interaction engagement, respect for cultural differences, interaction confidence, interaction enjoyment, and interaction attentiveness. Although this scale has been widely recognized and utilized by academicians across borders, Wang and Zhou (2016) critiqued its practicality in their study stated that some scholars perceived the IS of individuals as a developmental process; therefore, they preferred measuring their research participants' personal IS advancement multiple times. Nevertheless, the length and longevity of the IS scale proposed by Chen and Starosta (2000) obstruct its implementation when it is utilized with other qualitative and quantitative constructs. Moreover, the research participants may lose their concentration while engaging with long lists of scale items, and their honesty to respond to the questions may diminish owing to its long duration. To this end, Wang and Zhou (2016) initially broadened the content of Chen and Starosta's (2000) original IS scale by modifying 15 items from Barbaranelli, Caprara, Rabasca, and Pastorelli's (2003) Big Five Questionnaire to determine learners' demeanors encompassing their socialization, amiability, acceptance of novel experiences, awareness, and emotional balance and adapting 14 items from Wilson's (2013) Revised Sociocultural Adaptation Scale to identify their social adaptation to a foreign culture. Then, they shortened the measure following Stöber and Joormann's (2001) self-report abridgment procedures to prepare a scale in brevity. This modified scale contained five subscales contingent with Chen and Starosta's (2000) five factors interaction engagement, respect for cultural differences, interaction confidence, interaction enjoyment, and interaction attentiveness. And, each subscale encompassed three items. According to Wang and Zhou (2016), although the scale incorporated only 15 items, the performance of the new scale was estimated higher based on reliability, internal consistency, and validity compared to Chen and Starosta's (2000) original version.

To measure the IS in different cultural contexts with various research participant profiles, valid and reliable instruments are needed. For example, to measure IS of young learners, Pingault, Falissard, Côté, and Berthoz (2012) suggest that short scales are more appropriate to measure individual demeanors at multiple time zones over a certain period since longer scales may disturb younger subjects' concentration and challenge them to provide sincere responses for each item. Contingent with this statement, Wang and Zhou's (2016) IS-15 scale is singled out for the present research since it renders to assess younger learners' IS and attitudinal stance with ease without hampering the accuracy of the estimation.

Due to their cognitive, social, and emotional characteristics, young learners have different characteristics compared to adult learners. In this respect, conducting academic research necessitates the utilization of different methods and techniques (Tavil & Söylemez, 2020) to gather accurate data and produce reliable and valid results. Also, research related to children is challenging when compared to studies with adult language learners. In Aydın, Harputlu, Uştuk, Çelik and Güzel (2021), the underlying reasons for the challenging aspect of research with young language learners are discussed such as follows: (a) the lack of linguistic competence in the foreign language and even in their L1, (b) the ongoing and incomplete process of cognitive, psychological, and social development, (c) the need for extra visual, oral and physical interaction during data collection, (d) the lack of research methods and techniques specifically designed for young learners. Therefore, research with young language learners in the Turkish EFL context requires specific data collection instruments which are

developed and/or adapted to overcome the challenges listed above. From an intercultural awareness perspective, Zorba and Çakır (2019) also suggest further research on this topic with young learners.

Due to the increased interest in IS, scholars have investigated the IS in the Turkish context as well. The literature review on IS in Turkish and Turkish EFL context reveals that the studies used Chen and Starosta's (2000) scale and its adaptation into Turkish in their studies in descriptive studies mostly. With the Turkish adaptation of the scale, Üstün (2011) explored the IS levels of teacher candidates concerning various variables (such as their departments, size of hometown, and going abroad experiences). Particularly in the Turkish EFL context, there is a limited number of studies with a growing interest recently conducted with EFL learners and pre-service English teachers. Çiloğlan and Bardakçı (2019) investigated the relationship between IS and English language achievement and found a positive correlation. Also, Bal (2020) identified the level of EFL learners' IS and found that their IS level was above the average. In. a study by Altan (2018) in language teacher education context, it was found that ELT pre-service teachers had high respect for cultural differences. Recently, to meet the research needs for investigating intercultural awareness as a general concept at the secondary level, Zorba and Çakır (2019) developed a questionnaire which was in English, and IS was only one dimension of the scale.

1.2. The Present Study

The literature supports the idea that to be able to investigate the ICC and interculturality, IS is a fundamental concept to be explored and measured in second/foreign language learning contexts. The short version of the IS scale of Wang and Zhou (2016) was found both reliable and valid, and researchers have recently attempted to use the short IS scale to investigate the IS of the teachers and learners in the field of language education. However, they mainly use the original form of the scale. Since the language of the scale is important due to some factors such as learners' language proficiency and cross-cultural issues and elements which may differ in every educational context, the original version of the scale which is English will not provide reliable and valid results. Moreover, to investigate IS of young learners of English, the Turkish version of the scale needs more deliberate attempts such as simplification and moderation.

Previous studies (Altan, 2018; Çiftçi & Gürbüz, 2019; Harmandaoğlu, 2013; Kazazoğlu & Ece, 2021) focusing on ICC and IS in Turkey were mostly held with pre-service and novice English teachers by using Chen and Starosta's (2000) 's English questionnaire. There is a limited number of studies used the Turkish adaptation of the Chen and Starosta's (2000) scale (e.g., by Üstün (2011) with EFL learners (Bal, 2020; Çiloğlan & Bardakçı, 2019).

Even though the aforementioned scales are reliable and valid to be used with teenage and adult learners, there is still a need for a questionnaire suitable for the short-attention-span and limited linguistic skills of young learners. To the best knowledge of the researchers, there is no Turkish version of the IS scale of Wang and Zhou (2016) to use in Turkish contexts and specifically with young learners of English in Turkey. The length and linguistic intricacy of the existing IS scales may lead young learners to lose their concentration during their engagement with complex scale items and affect their integrity to provide honest responses to them. In this respect, it is necessary to offer an IS scale that is succinct and intelligible for young learners for more reliable and valid assessments. Moreover, to track the development of IS of EFL learners at any age group (with pre-and post-tests through some intercultural

interventions like telecollaboration) instead of a descriptive study with no intervention, researchers need a practical and reliable IS scale. Based on this background motivation and the research gap, the purpose of the study is two-fold: (a) to translate and validate the IS scale in Turkish EFL context, (b) to adapt and validate the Turkish version of the IS scale to be used for young English language learners in Turkish EFL context.

2. Method

Based on the study's two objectives, the methodology of the study was designed in two main stages by merging two studies. Each study consists of necessary and interrelational stages which must follow one another. This procedural approach was chosen to reach a more reliable and valid IS scale to be used for different contexts. Study 1 includes translation, implementation, and statistical analysis of the Turkish version while Study 2 includes simplification and moderation, implementation, and statistical analysis of the Turkish scale for young learners of English. The research methodology explained below was designed to provide two instruments to the researchers who might have different participant profiles.

2.1 Participants

For implementing the translated version of the scale (Study 1), the sample group was selected as 122 language learners studying in the faculties of Medical, Artificial Intelligence Engineering, and Aerospace Engineering of several universities all around Turkey. Convenience sampling was used due to reasons related to easy accessibility, geographical proximity, availability at a given time, or the willingness to participate are included for the study (Dörnyei, 2007). Of the participants, 70 (57.38%) were in their first year of college education and 52 (42.62%) were in their second year of university training. In this research group, 77 (63.11%) of the participants were female whereas 45 (36.89) of the participants were male. Moreover, their ages ranged between 18 and 21, and the mean age was computed as 19.22.

To implement the simplified version for young learners (Study 2), 95 young learners participated in the study to test the simplified version by the experts. Purposive sampling was used for the selection of the participants since the objective was to collect the data specifically from young foreign language learners. At this stage, the simplified version of the scale was tested with a Puppet Show role-play activity held by the drama expert. These young learners were 5th- and 6th--grade students whose ages ranged between 9 and 12. Later, to determine the reliability and validity values of the simplified scale, 115 young learners participated in the study. The sample group was detected as 74 (seventy-four) 5th-grade students with the 64.35% estimate and 41 (forty-one) 6th-grade students with the 47.83% estimate. Of the 5th and 6th grade learners, 4 (3.48%) were nine years old, 49 (42.60%) were ten years of, 54 (46.96%) were 11 years old, and 8 (6.96%) were twelve years old. Therefore, the mean age was computed as 10.57 for this analysis. Moreover, 70 (47.83%) of the participants were male whereas 55 (52.17%) of them were female in the research group.

2.1. Research Procedure and Data Analysis

The procedures for the adaptation of the scale for the Turkish version and young language learners Intercultural Sensitivity Scale developed by Wang and Zhou (2016) included the steps suggested by Erkuş and Selvi (2019) and Aydın et al. (2016). The steps followed were: (1) translation, (2) implementation, (3) statistical analysis, (4) simplification and moderation, (5) implementation, and (6) statistical analysis.

2.1.1. Step 1: Translation

First, Wenting Wang and Mingming Zhou, the developers of the scale to be used, were contacted via e-mail and requested written permission to adapt and use the scale for the current study. The researchers granted their consent for the present study. In addition, permission letters and approvals from the ethical committee were received from the Institute of Educational Sciences at Gazi University and the Ministry of National Education to implement the scale on secondary school students. When the consent and approval forms were delivered, the scale adopting procedures were followed to translate and adapt the scale. In the initial phase, five translators who were proficient in both English and Turkish and who had no familiarity with one another were selected to translate the original scale into Turkish in individual blind sessions. The first translator was a college instructor in the ELT department of a state university and a Ph.D. student in the ELT Department of another state university; therefore, he contributed to the translation process with his valuable academic knowledge and experience. The second translator was a foreign language teacher at a primary school and an MA diplomaed in the field of ELT. Hence, she had expertise in the spheres of foreign language learning and child development. The third translator was an American-born individual with full Turkish heritage and an MA student in the ELT department of a state university in Turkey. Her proficiency in both languages facilitated the translation process. The fourth translator was a BA graduate from the English Translation and Interpretation Department of a state university; thus, he contributed to the translation process with his educational and work experience. The last translator was an MA student in the Department of Psychological Counselling and Guidance of a state university. This translator had high English proficiency; therefore, she both made contributions to the interpretation process and functioned as the drama expert in the following moderation phase. The selected translators converted the original scale into English in independent blind sessions and sent them to the researcher via e-mail. The researcher combined the translated versions of the scale in a single Word document and held an online ZOOM meeting so that the translators could produce one interdependent scale out of these different versions. Since the IS-15 scale has clear statements, the translated works corresponded with one another. To this end, the meeting lasted only for 40 minutes. After the single translated version was generated with the consensus, the Turkish version of the scale was back-translated into English by two MA students in independent sessions. Then, these translators also organized an online meeting to unify their re-translated scales and produce a single back-translated version. In the end, the original and back-translated scales were compared by the researcher, and no significant difference was detected between them.

2.1.2. Step 2: Implementation

Once the crosscheck process ended in success, the English and Turkish versions of the scale were implemented over two time periods on 122 language learners studying at a university to detect the compatibility between them in terms of reliability and validity. For the administration phase, the sample group was selected. After the selection process, the original scale was employed through Google Forms, and the participants' responses were recorded to the SPSS software. Then, the Turkish version was implemented through the same platform after three weeks to reduce the recall of the previous answers. These answers were registered to the SPSS platform as well.

2.1.3. Step 3: Statistical Analysis

After implementing both English and Turkish versions of the IS scale, the collected data were statistically calculated to identify the similarities and dissimilarities between both scales. To measure the reliability, the internal consistency of both scales was estimated via Cronbach's Alpha measure. Then, the Kaiser-Meyer-Olkin (KMO) estimation was carried out to determine whether the sample size was appropriate for the factor analysis. Similarly, Bartlett's Test of Sphericity was conducted to find out if the data were normally distributed to begin the factor analysis. When the initial computations were completed, the exploratory and confirmatory factor analyses were regulated to verify the construct validity of both scales. In this process, exploratory factor analysis was executed with the principal component analysis and varimax rotation methods to identify the accords between the scale items. Furthermore, confirmatory factor analysis was performed to explore if there is a potent connection between the factors and if the factors define the framework effectively.

2.1.4. Step 4: Simplification and Moderation

After analysis on reliability and construct validity of the Turkish version of the scale, the next phase was determined as simplification and moderation of the scale for young learners. Although the language used in both versions was not difficult to comprehend, the scales were mainly designed for adult learners; therefore, they needed both simplifications to serve the linguistic and conceptual capacity of young learners and visual arrangement to be perceived as less threatening by them. For this purpose, the translators, who converted the original scale into Turkish, worked on the translated version this time to revise the language and make it more intelligible for children. They were asked to simplify the language of some of the items in blind sessions first. They were also requested to customize the appearance of the scale in line with the needs of young learners by replacing the points of the scale with emojis and changing the demographic information section to suit it for children. Then, they were invited to negotiate on their simplified versions in a ZOOM meeting and form a single joint simplified version. After the ultimate simplified version was produced in the meeting, the version was implemented first on 95 young learners by the drama expert through a Puppet Show role-play activity. In this process, the students were given a scenario such as having breakfast with a foreign friend or introducing themselves to a new foreign classmate. Then, they were requested to communicate with these foreigners effectively. The students not only performed but also watched their classmates' performances. After the showcase, the simplified version of the scale was distributed to the children so that they could answer the scale questions individually. In the responding process, the questions addressed and points not understood by the students were noted down by the researcher and drama expert to be evaluated in the next ZOOM meeting. The student performances and following scale completion procedures were recorded by them as well to be watched by all translators. The recordings and notes were sent to the translators via e-mail so that they could watch the videos, take their notes, and produce solutions to the encountered linguistic barriers.

One week later, a ZOOM session was organized by the researcher, and the translators reshaped the simplified scale following the students' needs. First, the word "peer", which was translated as "akran" in the Turkish version, was changed into the word "yaşıt" since the students couldn't grasp the former translated meaning. Second, the word "opinion", which was translated as "fikir" in the initial Turkish version, was replaced by the word "düşünce" since some students asked the meaning of the word "fikir", and they responded positively when it is explained as "düşünce" by one of their classmates. Since the students had difficulty

in understanding the word "narrow-minded", which was rendered into the word "dar görüşlü" in the initial Turkish version, they could not respond to the 6th item effectively. Therefore, the translators resorted to two Turkish teachers' assistance to enhance the comprehensibility of the item. Through the negotiations, the word was interpreted as "kendi düşüncelerinden başkasını düşünmemek" in the simplified Turkish version. Then, the phrase "being as social as possible", which was translated as "istediği kadar sosyal olmak" in the initial Turkish version was converted into "rahatça iletişim kurmak" in the simplified Turkish version. The reason behind the alteration was the fact that the students misinterpreted the meaning of the phrase as showing physical affection such as hugging and kissing; thus, reacted negatively due to the Covid-19 cases in their country. Finally, the wording in item 14 "being sensitive to subtle meanings", which was translated as "üstü kapalı ifadelere karşı duyarlı olmak" in the initial Turkish version, was transliterated as "açıkça söylenmeyen düşüncelere dikkat etmek" by the translators with the guidance of the Turkish teachers. The modification resulted from the fact that the students remained silent for a long time before responding to the item. In addition, they could not define whether being careful to subtle meanings is a positive quality when they are asked individually. Therefore, the amendment was seen as a necessity.

2.1.5. Step 5: Implementation

After the simplification and moderation process was finalized, the scale was piloted on 115 young learners. The implementation of this adapted version was conducted via a drama activity named "Film-Making". In this drama activity, the students were given some scenarios such as introducing themselves to a foreign classmate or asking a foreign classmate to play a game with them. Then, they were requested to communicate with these foreigners effectively. Students carried out this drama activity behind a TV screen made from cardboard. The students not only performed on this screen but also watched their classmates' performances. While watching their friends' acting, they were asked to use a remote controller and turn the volume up when they were unable to hear their friends' voices. After this showcase, the scales were distributed to the young learners so that they could answer the scale questions individually. In the responding process, the project teachers read aloud the questions one by one, and the young learners responded to the questions silently and individually. The rationale behind the utilization of the read-aloud technique was to make necessary explanations for the young learners when they remained unable to understand some scale items.

2.1.6. Step 6: Statistical Analysis

After implementation of the simplified scale, the data were analyzed to determine the reliability and validity values of the scale. After estimating the basic demographic measures, the same statistical analyses explained in Step 3 (Cronbach Alpha, the KMO and Bartlett's Test of Sphericity, exploratory and confirmatory factor analyses) were regulated on the SPSS platform.

3. Findings

3.1. Findings from Study 1

To determine the consistency between the original (English) and translated (Turkish) versions of the scale in this study, Cronbach's Alpha and Cronbach's Alpha Based on Standardized Items were reckoned. The Cronbach Alpha coefficients were estimated as 0.970

for the English version and as 0.973 for the Turkish version. Cronbach's Alpha coefficients extend between 0.00 and 1.00, and internal reliability escalates when the coefficients close into 1.00 value. The assessment criterion launched by George and Mallery (2003) highlight that the values between 0.90-1.00 are excellent. Based on the values and their interpretation above, the English and Turkish versions of the IS-15 scale were found to have correspondingly excellent internal consistency, which is given in Table 1 below.

Table 1.

Reliability Statistics

Versions	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Number of Items
English (Original)	0.970	0.970	15
Turkish (Translated)	0.973	0.973	15

Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) is utilized to reveal whether the sampling is suitable to carry out a factor analysis whereas Bartlett's Test of Sphericity correlates a study's estimated matrix to the identity matrix to check out whether the variables in both matrixes differ and function appropriately for the factor analysis. To conduct the exploratory factor analysis, KMO and Barlett's tests were regulated in the present study. The analyses, as given in Table 2, exhibited that the KMO value was 0.957 for the English version and 0.955 for the Turkish version. According to Tavsancil (2002), if the KMO value is between 0.70 and 0.79, the number of participants is good for analysis; if it is between 0.80 and 0.89, it is very good for analysis, and if it is between 0.90 and 1.00, it is perfect for analysis. Therefore, the estimated values higher than 0.90 in both scales pointed out that the sample size is perfect to continue with the factor analysis. In addition, the p-values were demonstrated as $x2=1890.881$ ($p< .00$) for the English version and $x2=2369.50$ ($p< .00$) for the Turkish version in Barlett's test. The p-values lower than 0.05 manifest the normal distribution of the data-set (Serçekuş, Vardar, & Başkale, 2020); therefore, the p-values of .000 indicated the appropriateness of the scales for the factor analysis.

Table 2.

KMO and Barlett's Test Computations

Versions	Kaiser-Meyer-Olkin Measure of Sampling Adequacy	Bartlett's Test of Sphericity	
English Version	0.957	Approx. Chi-Square	1890.881
		Df	105
		Sig.	.000
Turkish Version	0.955	Approx. Chi-Square	2369.500
		Df	105
		Sig.	.000

To determine the construct validity of the scale, exploratory factor analysis (EFA) was performed with principal component analysis and varimax rotation techniques. First, the item correlation matrixes of both scales were reviewed to examine the correspondence of the scale items with one another. The minimum correlation value of items was identified as 0.501 on the English scale and 0.573 on the Turkish scale. Since all correlation values were higher than 0.30 in both versions, no item was removed from the scales. Second, the explained variance of both scales and the scree plot test manifested a one-factor solution which signified the presence of only one prevalent mechanism in the whole structure. In

other words, the scales were determined as one dimensional with only one eigenvalue above the score of 1.00 (Figure 1 and 2). Since the Turkish version had a single factor solution as the original version, both scales demonstrated parallelism (Table 3). In addition, the total variance was explained as 71% for the English version and 72.86% for the Turkish version. According to Büyüköztürk (2010), the explained variance must be a minimum of 30% for one-dimensional scales. To this end, both the original scale and its translated version exposed good construct validity. However, the Turkish version displayed a slightly higher construct validity compared to the English version. Lastly, the factor loadings varied between 0.682 and 0.939 for the English scale and between 0.732 and 0.907 for the Turkish scale (Table 4). Factor loading scores over the value of 0.60 are considered as high factor loads; therefore, all items in both scales were regarded as effective.

Table 3.

Eigenvalues and Explained Variance of the English Version for Exploratory Factor Analysis

		Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
English Version	1	10.650	71.002	71.002	10.650	71.002	71.002
Turkish Version	1	10.930	72.867	72.867	10.930	72.867	72.867

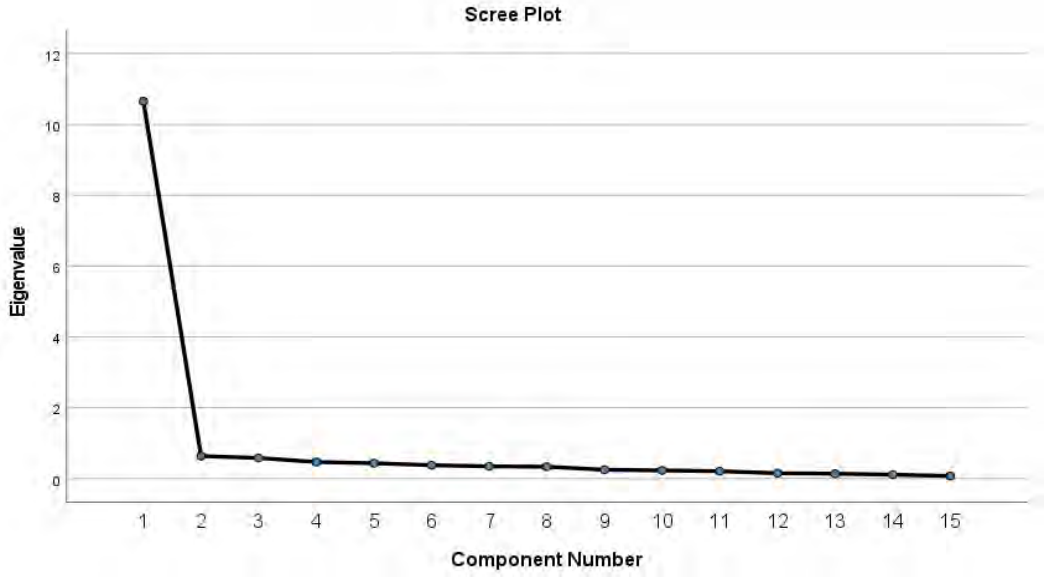

Figure 1. Component numbers and eigenvalues of the English version

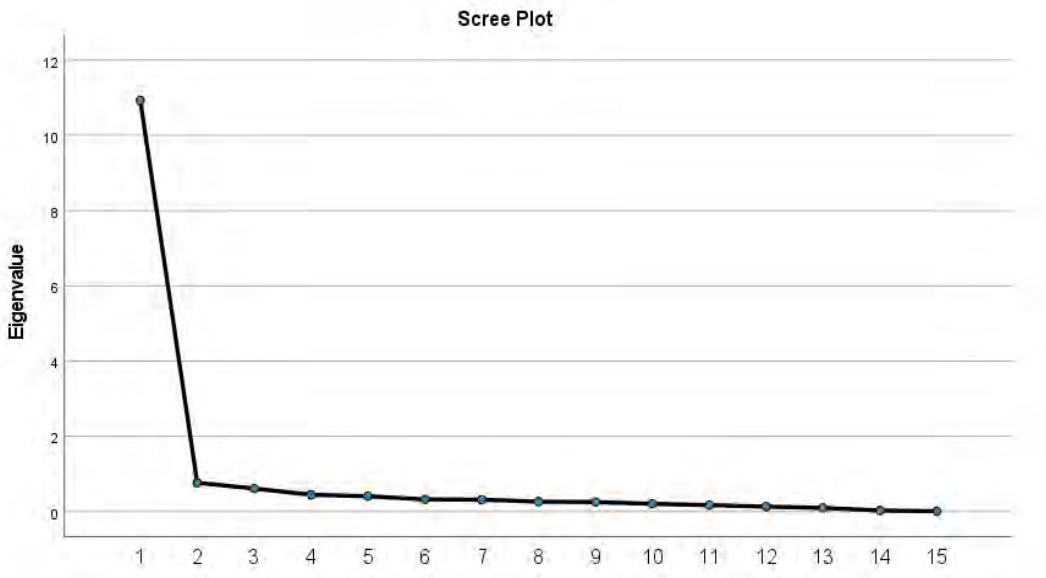

Figure 2. Component numbers and eigenvalues of the Turkish version

Table 4.

Factor Loading Values of the English and Turkish Versions

English Version		Turkish Version	
Items	Factor Loads	Items	Factor Loads
1	.939	1	.843
2	.855	2	.821
3	.883	3	.848
4	.779	4	.769
5	.682	5	.792
6	.771	6	.732
7	.868	7	.906
8	.896	8	.907
9	.897	9	.901
10	.817	10	.852
11	.863	11	.863
12	.875	12	.870
13	.848	13	.860
14	.830	14	.891
15	.803	15	.927

Confirmatory factor analysis (CFA), on the other hand, is another numerical procedure that is carried out to confirm the fundamental structure of measured variables. According to Yirci, Karakose, Uygun, and Ozdemir (2016), this analysis examines if there is a potent connection between measured variables and to what degree these variables define the whole structure. Therefore, it is used with EFA to determine the construct validity of assessment instruments. In the present study, CFA was regulated via the maximum likelihood estimation method, and the results were evaluated via model fit indices. The first fit index, Chi-square (x2), was used

to measure the correspondence between the hypothesized model and the computed variables. The estimation highlighted that both English and Turkish versions of the scale demonstrated a good fit for the model. The second index, degrees of freedom (df), computed the number of variables that demonstrated divergence. In this analysis, high degrees of freedom were detected for both versions, which signified the existence of a large sample size to continue the analysis. The proportion of the Chi-square to the degree of freedom was calculated as 1.828 for the original version and 1.28 for the Turkish version. Although the values were below the value of 3 in each scale and indicated a good fit, the Turkish version of the scale manifested a better model fit than the original version. The fourth index, the p-value was estimated as 0.00 in the English version and as 0.046 in the Turkish version. Even though the p-value in the original version was evaluated as a mediocre fit in the table, high scores obtained from the other model fit computations, including RMSEA, CFI, and GFI values altered its interpretation as a good fit in the current analysis. However, the Turkish version displayed an acceptable fit directly and showed improvement in the p-value compared to its English counterpart. The fifth index, the Root Mean Square Error of Approximation (RMSEA), was reckoned to identify the difference between the hypothesized and actual values. The reckoning reported that the English version presented an acceptable fit with a score of 0.08, whereas the Turkish version offered a good fit to the model with a score of 0.048. The sixth index, the Goodness of Fit Index (GFI) was computed to estimate the agreement between the hypothesized structure and actual covariance matrix, while the seventh index, the Adjusted Goodness of Fit Index (AGFI) was measured to rectify the GFI so that the complexity of the hidden variables could not adversely influence it. In the analysis, the GFI value was detected as 0.863 with the portrayal of acceptable model fit, and the AGFI value was established as 0.794 with the indication of mediocre fit for the English scale. On the other hand, the GFI measure was observed as 0.907 and interpreted as a good fit to the model, while the AGFI measure was revealed as 0.86 and labeled as an acceptable fit to the model for the Turkish scale. As the findings suggest, the translation process contributed to the amelioration of the model fit indices and yielded improved results. The Incremental Fit Indices (IFI) were measured to examine the contrast between the chi-square value for the hypothesized model and the null model. IFI includes two measurements labeled as Comparative Fit Index and Normed Fit Index. In this analysis, Comparative Fit Index (CFI) was computed to investigate the model fit by identifying the difference between the baseline model and the collected data. On the other hand, Normed Fit Index (NFI) was counted to find out the divergence between the chi-squared score of the hypothesized structure and the chi-squared score of the null model. The non-normed Fit Index (NNFI), which was also titled Tucker-Lewis Index (TLI), was regulated to overcome the negative biases, which could be acquired in the NFI estimation process. The aforementioned estimations highlighted that IFI, CFI, and NNFI values were over 0.95 in both English and Turkish scales; hence, the measures were interpreted as good fits to the model. Nevertheless, the NFI score was estimated as 0.926 in the English version and defined as an acceptable fit while it was assessed as 0.966 in the Turkish version and described as a good fit. Lastly, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), offer payments for good fits and penalties for escalating the number of predicted specifications (Eaton & Willoughby, 2018), were reckoned to compare both versions and identify the more effective one. Since the measurement tools with smaller AIC and BIC values represent higher validity, the Turkish version represented a better model fit in the analysis (Table 5).

Table 5.

Comparison of Model Fit Estimates between The English Version and the Turkish Version.

Measure	Scale Values		Good Fit	Acceptable fit	Fit (English Version)	Fit (Turkish Version)
	English Version	Turkish Version				
x2	146.215	102.425	$0 \leq x2 \leq 2df$	$2df < x2 \leq 3df$	Good Fit	Good Fit
Df	80	80		-	-	-
x2/df	1.828	1.28	≤ 3	≤ 5	Good Fit	Good Fit
P	0.00	0.046	$.05 < p \leq 1$	$.01 \leq p \leq .05$	Mediocre Fit	Acceptable Fit
RMSEA	0.08	0.048	≤ 0.05	0.05 - 0.08	Acceptable Fit	Good Fit
GFI	0.863	0.907	≥ 0.90	0.85 – 0.89	Acceptable Fit	Good Fit
AGFI	0.794	0.86	≥ 0.90	0.85 – 0.89	Mediocre Fit	Acceptable Fit
CFI	0.965	0.992	≥ 0.95	≥ 0.90	Good Fit	Good Fit
NFI	0.926	0.966	≥ 0.95	≥ 0.90	Acceptable Fit	Good Fit
NNFI / TLI	0.954	0.99	≥ 0.95	≥ 0.90	Good Fit	Good Fit
IFI	0.965	0.992	≥ 0.95	≥ 0.90	Good Fit	Good Fit
RMR	0.017	0.011	<0.05	<0.08	Good Fit	Good Fit
AIC	226.215	182.425	Smaller in comparison model		Good Fit	Better Fit
BIC	338.376	294.586	Smaller in comparison model		Good Fit	Better Fit

In conclusion, both the English scale and its Turkish version demonstrated high model fit indices in each measurement category. However, the Turkish scale exhibited more enhancement in model fit estimations in comparison to the original scale. Moreover, GFI and AGFI values are contingent on sample size (Hu & Bentler, 1995). Therefore, these indices decline when the model has a high intricate structure notably in small sample-sized examinations (Anderson & Gerbing, 1984). Conversely, RMSEA, NNFI, and CFI values are responsive to the model errors and not dependent on sample size (Hu & Bentler, 1998). In this regard, the x2/df, RMSEA, NNFI, and CFI values should be regarded first before the evaluation of GFI and AGFI values since they are not affected by the number of examinees in the analyses (Schermelleh-Engel, Moosbrugger, & Müller, 2003). Drawn on the information above, the mediocrity of the p and AGFI values in the English scale can be stated to be remedied by the other acceptable and good fit measures, including the x2/df, CFI, and NNFI scores.

3.2 Findings from Study 2

The Cronbach Alpha coefficients were reckoned as 0.94 for the simplified scale (Table 6), which was interpreted as an excellent internal consistency value according to the assessment protocol launched by George and Mallery (2003). Then, the KMO estimation was carried out to determine whether the sample size is appropriate for the factor analysis. Similarly, Bartlett's Test of Sphericity was conducted to find out if the data were normally distributed

to begin the factor analysis. The KMO value was calibrated as 0.924 in the computations, which was defined as perfect for analysis in line with Tavsancil's (2002) evaluation criteria (Table 7). On the other hand, the p-value was demonstrated as x2=1077.019 (p< .001) in Bartlett's tests, which indicated the normal distribution and fitness of the data-set to carry out factor analysis. When the KMO and Bartlett's Test scores verified the appropriateness of the simplified version for the factor analysis, EFA was regulated to detect the construct validity of the measurement tool.

Table 6.

Reliability Statistics of the Simplified Version

Versions	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Number of Items
Adapted Version	0.94	0.94	15

Table 7.

KMO and Barlett's Test Computations of the Adapted Version

Versions	Kaiser-Meyer-Olkin Measure of Sampling Adequacy	Bartlett's Test of Sphericity	
Adapted Version	0.924	Approx. Chi-Square	1077.019
		Df	105
		Sig.	<0.001

The EFA was performed with principal component analysis and varimax rotation techniques in this stage. First, the item correlation matrixes of both scales were reviewed to examine the correspondence of the scale items with one another. The minimum correlation value of items was identified as 0.498 on the simplified scale. Since all correlation values were higher than 0.30 in both versions, no item was removed from the scale. Second, the explained variance and the scree plot test manifested a one-factor solution which signified the presence of only one prevalent mechanism in the whole structure (Figure 3). In other words, the scale was identified as one dimensional with only one eigenvalue above the score of 1.00. In addition, the total variance was explained as 54.905% for this version (Table 8). According to Büyüköztürk (2010), the explained variance must be a minimum of 30% for one-dimensional scales. As a result, the simplified version exhibited good construct validity. Lastly, the factor loadings ranged between 0.653 and 0.838 for the scale. Factor loading scores over the value of 0.60 are considered as high factor loads; therefore, all items in this version were regarded as effective.

Table 8.

Eigenvalues and Explained Variance of the Adapted Version for Exploratory Factor Analysis

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.236	54.905	54.905	8.236	54.905	54.905

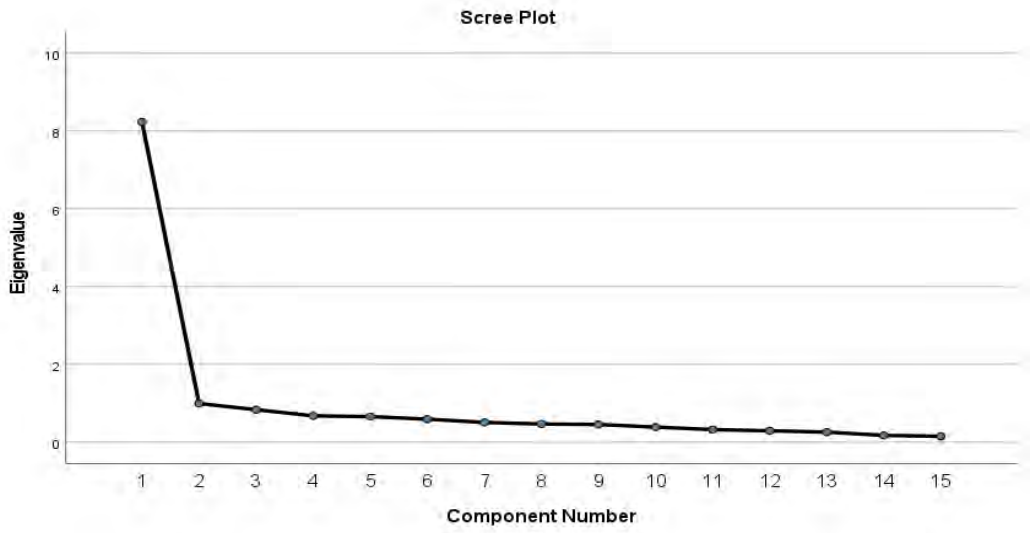

Figure 3. Component Numbers and Eigenvalues of the Adapted Version

Table 9.

Factor Loading Values of the Adapted Turkish Versions

Items	Factor Loads
1	.814
2	.770
3	.743
4	.719
5	.688
6	.736
7	.778
8	.805
9	.712
10	.669
11	.710
12	.838
13	.653
14	.702
15	.752

After the estimation of EFA values, CFA was carried out via the maximum likelihood measurement method, and the results were interpreted through model fit indices. The first fit index, Chi-square (x2) was used to measure the correspondence between the hypothesized model and the computed variables. The estimation highlighted that the simplified scale demonstrated a good fit for the model. The second index, degrees of freedom (df) computed the number of variables that demonstrated divergence. In this analysis, high degrees of freedom were detected for both versions, which signified the existence of a large sample size to continue the analysis. The proportion of the chi-square to the degree of freedom was calculated as 1.603 for the scale, which also indicated a good fit for the model. The fourth

index, the p-value was estimated as 0.00 in this version. Even though the p-value was labeled as a mediocre fit in the table below, high scores obtained from the other model fit computations including RMSEA, CFI, IFI, and GFI values altered its interpretation as a good fit in the current analysis. The fifth index, RMSEA was reckoned to identify the difference between hypothesized value and the actual value. The reckoning reported that the simplified version presented an acceptable fit with the score of 0.073. The sixth index, GFI was computed to estimate the agreement between the hypothesized structure and actual covariance matrix, while the seventh index, the AGFI was measured to rectify the GFI so that it could not be adversely influenced by the complexity of the hidden variables. In the analysis, the GFI value was detected as 0.879 with the portrayal of an acceptable model fit, and the AGFI value was established as 0.812 with the indication of a mediocre fit for this scale (Table 10). The eighth index, IFI was measured to examine the contrast between the chi-square value for the hypothesized model and the chi-square value for the null model. IFI includes two measurements labeled CFI and NFI. In this analysis, CFI was computed to investigate the model fit by identifying the difference between the baseline model and the collected data. On the other hand, NFI was counted to find out the divergence between the chi-squared score of the hypothesized structure and the chi-squared score of the null model. Moreover, the NNFI was regulated to overcome the negative biases, which could be acquired in the NFI estimation process. The aforementioned estimations highlighted that NFI and NNFI values were over 0.90 and interpreted as the acceptable model fits whereas the IFI and CFI values were over 0.95 and evaluated as good model fits. Lastly, the AIC and BIC values, which offer payments for good fits and penalties for escalating the number of predicted specifications (Eaton & Willoughby, 2018), were reckoned to compare the versions and identify the more effective one. Since the measurement tools with smaller AIC and BIC values represent higher validity, the simplified version represented a better model fit compared to the English version in the analysis.

Table 10.

Model Fit Estimates of the Adapted Version

Measure	Scale Values	Good Fit	Acceptable fit	Fit
x2	123.402	$0 \leq x2 \leq 2df$	$2df < x2 \leq 3df$	Good Fit
Df	77	-	-	-
x2/df	1.603	≤ 3	≤ 5	Good Fit
P	0.001	$.05 < p \leq 1$	$01 \leq p \leq .05$	Mediocre Fit
RMSEA	0.073	≤ 0.05	0.05 - 0.08	Acceptable Fit
GFI	0.879	≥ 0.90	0.85 – 0.89	Acceptable Fit
AGFI	0.812	≥ 0.90	0.85 – 0.89	Mediocre Fit
CFI	0.955	≥ 0.95	≥ 0.90	Good Fit
NFI	0.90	≥ 0.95	≥ 0.90	Acceptable Fit
NNFI / TLI	0.939	≥ 0.95	≥ 0.90	Acceptable Fit
IFI	0.956	≥ 0.95	≥ 0.90	Good Fit
RMR	0.62	<0.05	<0.08	Acceptable Fit
AIC	209.402	Smaller in comparison model		Better Fit Compared to the English Version
BIC	327.434			

By taking all model fit indices and interpretations into consideration, it can be concluded that the CFA analysis revealed overall acceptability and appropriateness in each measurement category since the mediocre fitness of the p and AGFI values were enhanced by the remaining high fit indices in the analysis. The rationale lies in sampling dependency of the

GFI and AGFI values. According to Hu and Bentler (1995), GFI and AGFI values depend on sample size. Hence, these indices diminish when the model has a highly complex structure, notably in small sample-sized examinations (Anderson & Gerbing, 1984). Conversely, RMSEA, NNFI, and CFI values are responsive to the model errors and not dependent on sample size (Hu & Bentler, 1998). In this regard, the x2/df, RMSEA, NNFI, and CFI values are regarded first before the evaluation of GFI and AGFI values since they are not affected by the number of examinees in the analyses (Schermelleh-Engel et al., 2003). Drawn on the information above, the mediocrity of the p and AGFI values in the simplified scale were remedied by the other acceptable and good fit measures, including the x2/df, CFI, and NNFI scores.

4. Discussion and Conclusion

This study aimed to translate and validate the IS scale in the Turkish EFL context and adapt and validate the Turkish version of the IS scale to be used for young English language learners in the Turkish EFL context. To these ends, the study included two main stages which include the translation and implementation of the IS scale into Turkish first and then simplification, moderation, and the implementation of the Turkish version of the IS scale. The findings from the first stage revealed that the IS scale in the Turkish language has correspondingly excellent internal consistency. Also, the Turkish version displayed a slightly higher construct validity compared to the English version. Both the English scale and its Turkish version demonstrated high model fit indices in each measurement category. However, the Turkish scale exhibited more enhancement in model fit estimations in comparison to the original scale. As a result, it can be stated that the Turkish version of IS scale is an appropriate instrument to measure the IS levels of foreign language learners in Turkish EFL contexts (See Appendix A).

Furthermore, the findings from the second stage to determine the validity of the IS scale for young language learners also indicate similar reliability and validity results. The simplified version of the Turkish IS scale also has an excellent internal consistency value and good construct validity. The findings related to the factor loadings indicate high factor loads; thus, all items in the simplified version were regarded as effective. Therefore, it can be claimed that the simplified version of the Turkish IS scale has the potential to use as an appropriate measurement tool for IS levels to young language learners in Turkish EFL contexts (see Appendix B). Also, the implementation of these stages as the methodology of the study has contributed to the field by providing a sample for future studies. Such methodology is needed for the studies on the adaptation and validation process of scales into Turkish and young learners.

Based on the results, the current study provides two reliable and valid instruments which can be used in different educational contexts and with different participant groups. The Turkish IS scale can be used with adult learners and also teachers for different purposes in any educational context: (a) to identify their current IS level for understanding their capacity to engage in intercultural encounters; (b) to measure the IS development of learners who are participating in exchange programs (etwinning, Erasmus+, virtual exchange) to be implemented before and after the exchanges; (c) to measure the IS development of foreign language learners with the ICC; (d) to provide quantitative data which will be followed by qualitative data such as reflective journals and interviews during a research intervention.

Furthermore, the Turkish scale for young learners provides several opportunities for research in the field. Since previous studies on IS and ICC are mostly conducted at the tertiary level, there is a need for research in primary and secondary educational contexts. In foreign language teaching, the ICC development and the IS need to be studied at lower levels and with young learners to identify their needs for further language teaching practices. Therefore, the Turkish scale for young learners can be used to identify the current level of IS of young learners which will be promising for language teachers and researchers to integrate more intercultural activities into their language classrooms. The findings on IS will provide insights for developing materials and activities to enhance their IS and ICC.

There were some limitations to this study. The study is limited to the young learners from the 5th and 6th grades in a secondary school EFL context. Young learners in lower grades were not included in the study. The simplification process of the IS scale through role-playing and the puppet show was also limited due to the classroom time allocated for English courses.

Note on Ethical Issues

The authors confirm that ethical approval was obtained from Gazi University (Approval Date: 06 /05 /2021).

References

Alaei, M. M., & Nosrati, F. (2018). Research into EFL teachers' intercultural communicative competence and intercultural sensitivity. *Journal of Intercultural Communication Research, 47*(2), 73-86.

Altan, M. Z. (2018). Intercultural sensitivity. *Journal of Intercultural Communication, 46*(1), 1-17.

Anderson, J. C., & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika, 49*(2), 155–173.

Aydın, S., Harputlu, L., Savran Çelik, S., Uştuk, Ö., Güzel, S., & Genç, D. (2016). Adapting scale for children: A practical model for researchers. *International Contemporary Educational Research Congress,* Muğla, Turkey. https://doi.org/10.14527/9786053186380

Aydın, S., Harputlu, L., Uştuk, Ö., Çelik, Ş. S., & Güzel, S. (2021). Difficulties in collecting data from children aged 7–12. *International Journal of Teacher Education and Professional Development (IJTEPD), 4*(1), 89-101.

Bal, N. G. (2020). Intercultural sensitivity of EFL learners at a state university. Journal of *Language Teaching and Learning, 10*(2), 1-19.

Barbaranelli, C., Caprara, G. V., Rabasca, A., & Pastorelli, C. (2003). A questionnaire for measuring the Big Five in late childhood. *Personality and Individual Differences, 34*(4), 645–664.

Bennett, J. M. (2011). Developing intercultural competence for international education faculty and staff. In *2011 AIEA Conference Workshop*, San Francisco, CA. (Vol. 25, p. 2018).

Bennett, M. J. (1986). A developmental approach to training for intercultural sensitivity. *International Journal of Intercultural Relations, 10*(2), 179–196. https://doi.org/10.1016/0147-1767(86)90005-2

Bennett, M. J. (2004). Becoming interculturally competent. In J.S. Wurzel (Ed.), *Toward multiculturalism: A reader in multicultural education* (pp. 62-77). Newton, MA: Intercultural Resource Corporation.

Berry, J. W., Poortinga, Y. H., Breugelmans S. M., Chasiotis A., & Sam D. L. (2011). *Cross-cultural psychology: Research and applications.* Cambridge: Cambridge University Press.

Büyüköztürk, S. (2010). *Data analysis handbook for social sciences.* Ankara: Pegem Academy.

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

Chen, G. M., & Starosta, W. J. (1996). Intercultural communication competence: A synthesis. *Communication Yearbook, 19*, 353-383.

Chen, G. M., & Starosta, W. J. (1998). *A review of the concept of intercultural awareness.* Communication Studies Faculty Publications.

Chen, G. M., & Starosta, W. J. (2000). The development and validation of the intercultural sensitivity scale. *Human Communication, 3,* 1–15.

Çiftçi, E. Y., & Gürbüz, N. (2019). Intercultural sensitivity orientations before short-term study abroad: A qualitative study on prospective English language teachers. *The Qualitative Report, 24*(6), 1319-1337.

Çiftçi, E. Y., Karaman, A. C., & Daloğlu, A. (2020). 'No one is superior to another: Tracing intercultural development in a year-long study abroad programme. *The Language Learning Journal,* 1–13. https://doi.org/10.1080/09571736.2020.1802770

Çiloğlan, F., & Bardakçı, M. (2019). The Relationship between Intercultural Sensitivity and English Language Achievement. *Journal of Language and Linguistic Studies, 15*(3), 1204-1214.

Deardorff, D. K. (2006). Identification and Assessment of Intercultural Competence as a Student Outcome of Internationalization. *Journal Studies in International Education, 10*(3), 241-266.

Dong, Q., Day, D. K., & Collaço, M. C. (2008). Overcoming ethnocentrism through developing intercultural communication sensitivity and multiculturalism. *Human Communication, 11*(1), 27– 38.

Douglas, S. R., & Rosvold, M. (2018). Intercultural Communicative Competence and English for Academic Purposes: A Synthesis Review of the Scholarly Literature. *Canadian Journal of Applied Linguistics Articles, 21*(1), 23–42.

Dörnyei, Z. (2007). *Research methods in applied linguistics.* New York: Oxford University Press.

Eaton, P., & Willoughby, S. D. (2018). Confirmatory factor analysis applied to the Force Concept Inventory. *Physical Review Physics Education Research, 14*(1). https://doi.org/10.1103/physrevphyseducres.14.010124

Erkuş, A., & Selvi, H. (2019). *Psikolojide ölçme ve ölçek geliştirme III: Ölçek uyarlama ve norm geliştirme.* Pegem Akademi, Ankara.

Garrett-Rucks, P. (2014). Measuring instructed language learners' IC development: Discrepancies between assessment models by Byram and Bennett. *International Journal of Intercultural Relations, 41*, 181–191.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference 18.0 update (11th ed.).* Boston: Allyn & Bacon/Pearson.

Hammer, M. R., Bennett, M. J., & Wiseman, R. (2003). Measuring intercultural sensitivity: *The intercultural development inventory. International Journal of Intercultural Relations, 27*(4), 421–443.

Harmandaoğlu, E. (2013). *Promoting intercultural competence of teacher trainees in English language departments via web 2.0 tools: A case study* (Unpublished master's thesis). Gazi University, Ankara, Turkey.

Hu, L., & Bentler, P. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling. Concepts, issues, and applications* (pp. 76−99). London: Sage.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure analysis: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424−453.

Kazazoğlu, S., & Ece, E. (2021). Exploring pre-service and novice EFL teachers' intercultural sensitivity. *Novitas-ROYAL (Research on Youth and Language), 15*(1), 76-89.

Munezane, Y. (2019). A new model of intercultural communicative competence: bridging language classrooms and intercultural communicative contexts. *Studies in Higher Education*, 1-18.

Peng, S. Y. (2006). A comparative perspective of intercultural sensitivity between college students and multinational employees in China. *Multicultural perspectives, 8*(3), 38-45.

Pingault, J. B., Falissard, B., Côté, S., & Berthoz, S. (2012). A new approach of personality and psychiatric disorders: a short version of the affective neuroscience personality scales. *Public Library Of Science, 7*(7), e41489.

Rantz, F., & Horan, P. (2005). Exploring intercultural awareness in the primary modern language classroom: The potential of the new model of European language portfolio developed by the Irish modern languages in primary schools initiative (MLPSI). *Language and Intercultural Communication, 5*(3–4), 209–221. https://doi.org/10.1080/14708470508668896

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8(*2), 23–74.

Serçekuş, P., Vardar, O., & Başkale, H. (2020). The validity and reliability study of the Turkish version of Assessment of Survivor Concerns (ASC) Scale. *Journal of Psychiatric Nursing, 11*(3), 220–227. https://doi.org/10.14744/phd.2020.63308

Sinicrope, C., Norris, J., & Watanabe, Y. (2007). Understanding and assessing intercultural competence: A summary of theory, research, and practice (Technical report for the foreign language program evaluation project). *Studies in Second Language Acquisition, 26*(1).

Stöber, J., & Joormann, J. (2001). A short form of the Worry Domains Questionnaire: Construction and factorial validation. *Personality and Individual Differences, 31*, 119–126.

Tavil, M., & Söylemez, S. (2020). How to do research with young and very young learners? In H. H. Uysal (Ed.), *Political, pedagogical and research insights into early language education (pp.* 82–94). UK: Cambridge Scholars Publishing.

Tavsancil, E. (2002). *Measurement of attitudes and data analysis with SPSS*. Ankara: Nobel Publishing.

Üstün, E. (2011). *Öğretmen Adaylarının Kültürlerarası Duyarlılık ve Etni kmerkezcilik Düzeylerini Etkileyen Etmenler*. Yayınlanmamış Yüksek Lisans Tezi. Yıldız Teknik Üniversitesi.

Wang, W., & Zhou, M. (2016). Validation of the short form of the Intercultural Sensitivity Scale (ISS-15). *International Journal of Intercultural Relations, 55*, 1–7. https://doi.org/10.1016/j.ijintrel.2016.08.002

Wilson, J. (2013). *Exploring the past, present, and future of cultural competency research: the revision and expansion of the sociocultural adaptation construct*. (Unpublished Doctoral Dissertation). Victoria: University of Wellington.

Yirci, R., Karakose, T., Uygun, H., & Ozdemir, T. Y. (2016). Turkish adaptation of the Mentorship Effectiveness Scale: A validity and reliability study. *EURASIA Journal of Mathematics, Science and Technology Education, 12*(4). https://doi.org/10.12973/eurasia.2016.1440a

Zorba, M. G., & Çakır, A. (2019). A case study on intercultural awareness of lower secondary school students in Turkey. *Novitas-ROYAL, 13*(1), 62–83.

Appendix A. Kültürlerarası Duyarlılık Ölçeği (Wang & Zhou, 2016) – Türkçe Versiyonu

Lütfen her ifade için "kesinlikle katılıyorum" için, 5'i, "katılıyorum" için 4'ü, "kısmen katılıyorum" için 3'ü, "katılmıyorum" için 2'yi ve "kesinlikle katılmıyorum" için 1'i işaretleyiniz.

	1	2	3	4	5
1. Farklı kültürlerden insanlarla iletişim kurmaktan keyif alırım.					
2. İletişimimiz boyunca, farklı kültürden olan akranıma sık sık olumlu cevaplar veririm.					
3. Farklı kültürlerden insanlarla konuşmam gereken durumlardan kaçınırım.					
4. Farklı kültürlerden insanlarla bir arada olmaktan hoşlanmam					
5. Farklı kültürlerden insanların fikirlerini kabul etmem.					
6. Diğer kültürlerden insanların dar görüşlü olduğunu düşünüyorum					
7. Farklı kültürden insanlarla iletişim kurma konusunda kendimden oldukça eminim.					
8. Farklı kültürlerden insanlarla iletişim kurarken kendime güvenirim					
9. Farklı kültürlerden insanlarla iletişim kurarken istediğim kadar sosyal olabilirim.					
10. Farklı kültürden insanlarla iletişim kurarken sıklıkla kendimi işe yaramaz hissederim.					
11. Farklı kültürden insanlarla iletişim kurarken kolayca üzülürüm					
12. Farklı kültürden insanlarla birlikteyken sık sık cesaretim kırılır					
13. Farklı kültürlerden insanlarla etkileşim kurarken çok dikkatliyimdir					
14. İletişimimiz sırasında, farklı kültürden olan akranımın üstü kapalı ifadelerine karşı duyarlıyım.					
15. Farklı kültürlerden insanlarla etkileşim kurarken mümkün olduğu kadar fazla bilgi edinmeye çalışırım					

Appendix B. Kültürlerarası Duyarlılık Ölçeği (Wang & Zhou, 2016) – Türkçe ve Çocuklar için Adapte Edilmiş Versiyonu

Lütfen her cümle için "kesinlikle evet" için  *, " evet" için*  *"kısmen evet" için*  *"hayır" için*  *ve "kesinlikle hayır" için*  *işaretleyiniz.*

					
1. Farklı kültürlerden insanlarla iletişim kurmaktan keyif alırım.					
2. İletişimimiz boyunca, farklı kültürden olan yaşıtıma sık sık olumlu cevaplar veririm.					
3. Farklı kültürlerden insanlarla konuşmam gereken durumlardan uzak dururum.					
4. Farklı kültürlerden insanlarla bir arada olmaktan hoşlanmam					
5. Farklı kültürlerden insanların düşüncelerini kabul etmem.					
6. Diğer kültürlerden insanların kendi düşüncelerinden başkasını kabul etmediğini düşünüyorum.					
7. Farklı kültürden insanlarla iletişim kurma konusunda kendimden oldukça eminim.					
8. Farklı kültürlerden insanlarla iletişim kurarken kendime güvenirim					
9. Farklı kültürlerden insanlarla rahatça iletişim kurabilirim.					
10. Farklı kültürden insanlarla iletişim kurarken sıklıkla kendimi işe yaramaz hissederim.					
11. Farklı kültürden insanlarla iletişim kurarken kolayca üzülürüm					
12. Farklı kültürden insanlarla birlikteyken sık sık cesaretim kırılır					
13. Farklı kültürlerden insanlarla iletişim kurarken çok dikkatliyimdir					
14. İletişimimiz sırasında, farklı kültürden olan yaşıtımın açıkça söylemediği düşüncelere karşı dikkatliyimdir.					
15. Farklı kültürlerden insanlarla iletişim kurarken mümkün olduğu kadar fazla bilgi edinmeye çalışırım					