

Validation of Persian Translation of the Intercultural Sensitivity Scale Using Rasch Model

Elahe Moradi

Ferdowsi University, Mashhad, Iran

Abstract: This study subjected the Intercultural Sensitivity scale consisting of 24 items to Rasch analysis in a sample of 200 participants. The scale was translated into Persian and administered to graduate and undergraduate university students. Analysis of data showed that there were no misfitting items. Furthermore, no item manifested gender differential item functioning (DIF). All the thresholds were ordered and respondents could distinguish well between categories of the scale. The results of principal component analysis (PCA) of standardized residuals revealed that there were two contrasts with eigenvalues above two. Deleting positively and negatively loading items separately did not improve model fit. Thus, the content of items was investigated and it became clear that most of the positively loading items covered those items which have negative content and conversely, negatively loading items encompassed the items which have positive content. This brings about two psychometric dimensions in this scale. 1) Wording of the items revealed that the existence of items with negative wording in the scale results in statistical artifacts, and 2) The secondary dimension here could be interpreted to be an artifact of the wording. It was concluded that the scale is unidimensional and enjoys acceptable psychometric properties.

Keywords: intercultural sensitivity, Rasch Rating Scale model, unidimensionality, validity.

Intercultural Sensitivity and its Assessment

Intercultural sensitivity is considered one of the main aspects of intercultural communication competence which is gaining attention in various fields. Chi and Suthers (2015) focused on the achievement of intercultural communication competence and cultural information through the existence of close ties and collaborations among members of different cultures. It is essential to recognize differences among cultures. Also, any prejudices that impede individuals from communicating effectively with people from different cultures should be identified. Prejudice is generally referred to as the unfair, biased, and intolerant attitudes towards people from different cultures because they belong to various religions, races, or nationalities. (Samovar & Porter, 1991). Thus, a skilled communicator must be aware of such barriers and attempt to reduce them by continually checking understanding and offering appropriate feedback to their culturally-distinct counterparts.

Nowadays the tendency towards globalization has increased the importance of being competent in communicating with people of different cultural backgrounds. Intercultural communication is the communication between people from different cultures, and good

intercultural communication requires an interculturally sensitive attitude and cultural knowledge. Some authors (Flores & Palacios, 2018) defined interculturality as “the social capacity of people to build their future, taking advantage of the lessons of their historical experience and the real and potential resources of their culture.”

According to Chen and Starosta (2000) intercultural sensitivity is one of the prerequisites of intercultural communication competence, and people with higher levels of intercultural sensitivity will be more confident in communication as their understanding of cultural differences becomes more complex.

Intercultural sensitivity refers to the affective dimension of intercultural communication competence. Affective dimension refers to the emotions that may motivate a person to engage in intercultural interactions. (Chen & Starosta, 1996). This dimension focuses on the feelings and emotions created by specific situations, persons, and environments that are not the same as one’s own culture. (Carothers, 2018; Halpern, 2018; Moradi & Ghabanchi, 2019).

Bhawuk (1992) measured intercultural sensitivity by investigating the various ways people can negotiate based on their cultures, open-mindedness towards other cultures, and flexibility of the behavior in unfamiliar ways which is determined by the norms of other cultures. Bhawuk and Brislin (1992) suggested that intercultural sensitivity is a feedback an individual gives to people from other cultures. They stated that:

To be effective in other cultures, people must be interested in other cultures, be sensitive enough to notice cultural differences and also be willing to modify their behavior as an indication of respect for people of other cultures. A reasonable term that summarizes these qualities of people is intercultural sensitivity. (p. 416)

In order to engage in intercultural interactions, one must be open to an idea of others and otherness. This is why it is crucial that individuals should be educated about intercultural values and become competent in intercultural interactions. (Delante, 2020; Tétreault et al. ,2020; Vu, 2019).

Research on intercultural communication competence has mainly attempted to produce models based on individual traits that relate individual attitudes and skills to some measure of interculturally successful behaviors. Among studies in this line of research, Chen and Starosta's (1996) model of intercultural communication competence draws much special interest. The model is comprised of three conceptual dimensions of intercultural communication competence, including intercultural awareness, intercultural sensitivity, and intercultural adroitness. Based on this model, Chen and Starosta (2000) further explained the nature and constituents of intercultural sensitivity and developed an instrument to measure the concept. They developed a new intercultural sensitivity measurement scale to “integrate features of both cross-cultural attitude and behavioral skills models” (Fritz et al., 2001, p. 54). This scale is a 24-item questionnaire aimed at measuring intercultural sensitivity. The scale includes five factors on which its items are based: (1) interaction engagement, (2) respect for cultural differences, (3) interaction confidence, (4) interaction engagement, and (5) interaction attentiveness (Chen & Starosta, 2000, p. 98).

In developing and validating Intercultural Sensitivity Scale (ISS), Chen and Starosta (2000) conducted three studies; the first one was to evaluate the dimensionality of the items using Exploratory Factor Analysis (EFA); the second was undertaken to examine the concurrent validity of the ISS, and the third one was to investigate the predictive validity of the scale. The Cronbach’s alpha coefficients for the first two studies were reported 0.86 and 0.88 respectively. “ISS has demonstrated strong reliability and appropriate concurrent and predictive validity. While further

research is needed to replicate the properties of the ISS, the scale shows promise for use as a measure of intercultural sensitivity.” (Chen & Starosta, 2000).

Several other studies have been done to scrutinize the validity and reliability of the Intercultural Sensitivity Scale (ISS). Fritz et al. (2001) used the ISS on German students in Germany, to see if the scale could be utilized among different cultural groups. According to Fritz et al. (2001) “The results of confirmatory factor analysis in this study by using a German sample confirmed the validity of the overall structure of Chen and Starosta’s instrument on the measurement of intercultural sensitivity” (p. 57).

In the same line, Peng et al. (2005) measured the intercultural sensitivity levels of Thai and Chinese nationals. They used the scale and reported it as a reliable one, but no information was provided on the reliability of the scale. Peng (2006) employed ISS in China and reported the scale to be reliable with an overall Cronbach’s alpha of 0.8. Peng et al. (2005) and Peng (2006) did not examine the validity of the scale. Also Dong et al. (2008) found ISS reliable, but they did not investigate the validity of Chen and Starosta’s Intercultural Sensitivity model. Since the number of validation studies related to ISS is limited, there is an ongoing discussion among scholars whether ISS is a generic or culture-free model even though it was validated with American and German students (Chen & Starosta, 2000; Fritz et al., 2005; West, 2009). Chen and Starosta’s model of IS demands further validation testing, particularly for non-western countries. Therefore, the current study employed the Intercultural Sensitivity Scale (ISS), the measurement scale which has been designed by Chen and Starosta (2000) for further validation. This model was selected because it takes into account all the dimensions of intercultural communication competence, and has specific intentions to measure intercultural sensitivity of people from various cultural backgrounds. The objective of this study was to examine the psychometric qualities of the Intercultural Sensitivity Scale using Rasch Rating Scale model to determine whether the Persian translation of the scale confirmed the validity of it. The purpose of the present study is to extend previous works on ISS by validating the instrument within another cultural setting.

Methodology

Participants

The Persian translation of Intercultural Sensitivity Scale was administered to 200 Iranian graduate and undergraduate students at two universities in Mashhad, Iran. ($Mean_{age}=24.79$ and $SD=5.67$). The participants were selected based on availability sampling. The sample included 68 male and 132 female participants.

Research Tool

Intercultural Sensitivity Scale designed by Chen and Starosta (2000) is composed of 24 items using a 5-point Likert scale ($1=strongly\ disagree$, $2=disagree$, $3=uncertain$, $4=agree$, $5=strongly\ agree$). Items 2, 4, 7, 9, 12, 15, 18, 20 and 22 require reverse coding. To conduct this study, the scale was translated from English into Persian. The Persian translation was back-translated into English by two experienced translators and finally it was revised based on their comments.

Analysis and Results

The Andrich (1978) Rating Scale Model (RSM) was used to examine the psychometric qualities of the Intercultural Sensitivity Scale. The data were analyzed using the computer program WINSTEPS Rasch software, version 3.72.3. (Linacre, 2006).

The Rasch model is a latent variable model and observed variables are related to a latent structure. If the data fit the Rasch model it is evidence that there is a latent trait underlying the observed responses and there is monotone increasing relationship between them (Baghaei & Shoahosseini, 2019). These two conditions are necessary for validity (Borsboom et al., 2004). One of the most important and robust requirement of the Rasch model is the unidimensionality requirement. This assumption implies that a test should measure one single construct or dimension at a time and one of the main properties of this model and its superiority over true score theory is the invariance of item and person measures.

Examining the infit and outfit mean-squares associated with the responses to each item of the scale indicated that there are no misfitting items. Values above 1.5 are problematic. (Linacre, 2009). Misfitting items are indications of construct-irrelevant variance (Baghaei, 2008). As is evident from Table 1, all infit and outfit values are within the acceptable limit.

Table 1
Item Measures and Fit Statistics for Intercultural Sensitivity Scale

Item	Measure	Error	Infit MNSQ	Outfit MNSQ
1	-0.69	0.10	0.98	0.94
2	-0.98	0.11	1.45	1.30
3	0.44	0.08	0.64	0.64
4	0.95	0.08	1.02	1.02
5	1.12	0.07	0.77	0.79
6	0.34	0.08	0.92	0.93
7	-0.43	0.09	1.11	1.00
8	-1.42	0.12	0.80	0.78
9	-0.39	0.09	1.06	0.98
10	0.30	0.08	0.75	0.74
11	-0.56	0.09	0.90	0.95
12	0.21	0.08	1.27	1.30
13	0.37	0.08	1.19	1.30
14	0.49	0.08	0.79	0.80
15	-0.20	0.09	1.33	1.42
16	-1.02	0.11	1.02	0.93
17	-0.48	0.09	0.93	0.88
18	-0.22	0.09	1.10	1.12
19	0.28	0.08	1.14	1.25
20	0.29	0.08	1.36	1.41
21	1.09	0.07	1.29	1.34
22	-0.11	0.08	1.01	0.95
23	0.35	0.08	0.87	0.89
24	0.27	0.08	0.98	0.96

The Separation coefficient is an index that indicates the sensitivity of a scale in revealing differences in person measures. Person separation coefficient shows the number of statistically

separate ability strata that a scale or test can distinguish. This provides evidence for the external validity of the scale or test. (Wolf & Smith, 2007).

Low person separation is defined as separation < 2 , and low person reliability is considered as reliability < 0.8 (Linacre, 2009). Person separation index indicates whether the instrument is sensitive enough to distinguish between high ability and low ability participants. Findings revealed a separation value of 2.09 which is acceptable for group comparisons, but not individual comparisons and it can be concluded that intercultural sensitivity scale distinguishes well between participants with high and low cultural sensitivity.

The value of person reliability is 0.81 (> 0.8). This indicates that the ordering of persons can be marginally replicated if the sample is given the set of parallel items that measure the same construct.

Item separation indicates the number of different ability or difficulty strata that the test can identify. Item separation values < 3 , and item reliability values < 0.9 are considered as low item separation and reliability (Linacre, 2009). The separation value of the scale is 7 which is an acceptable index. Thus, the representativeness of the items of the intercultural sensitivity scale is confirmed. Also, item reliability was estimated as 0.98 (> 0.9), which can be considered an acceptable value.

Thresholds identify the points on the rating scale where the probability of being observed in either of two adjacent categories is equal. Disordered Andrich thresholds show that some categories on the latent variable are narrow. Disordered thresholds do not violate Rasch models, but they influence the interpretation of how the rating scale functions. (Linacre, 1999).

Table 2
Category Statistics for Intercultural Sensitivity Scale

Category	Count (%)	Observed average	Infit MNSQ	Outfit MNSQ	Threshold
1	115(2)	0.37	1.61	1.90	-
2	425(9)	0.25	1.11	1.16	-1.32
3	1122(23)	0.52	0.87	0.80	-0.58
4	1562(33)	1.00	1.08	0.96	0.52
5	1576(33)	1.81	0.83	0.88	1.39

As is shown in Table 2, all the thresholds are ordered and there is no need to collapse rating scale categories. This indicates that the respondents could distinguish among categories well.

Gender DIF

Lack of invariance of item parameters across different samples is referred to as differential item functioning (DIF). (Lord, 1980). DIF is regarded as evidence of item bias and DIF measure shows the difficulty of each item of the scale for each group of male and female respondents. According to Baghaei and Cassady (2014),

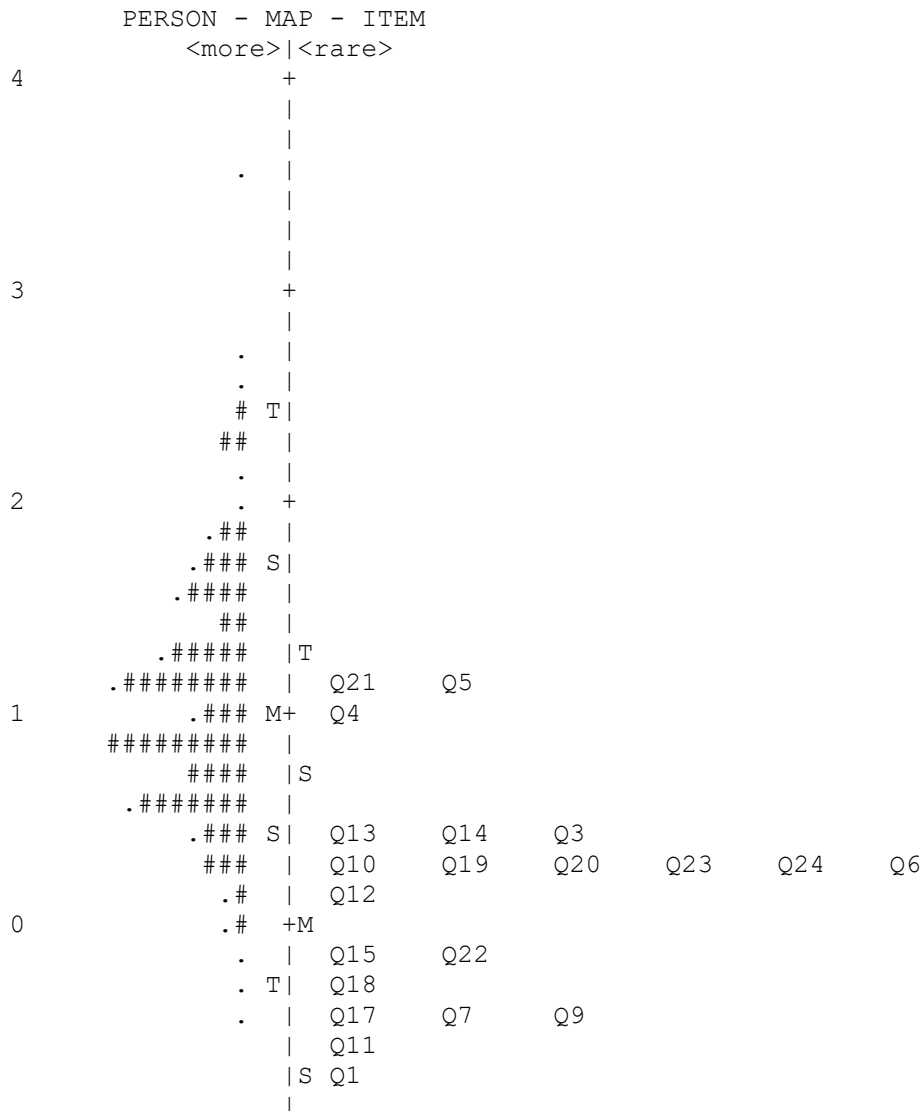
DIF occurs for an item when respondents with the same location on the latent trait have different probabilities of endorsing the item. DIF is an indication of the change of construct for different subpopulations of respondents and is a serious threat to test validity. (p. 7)

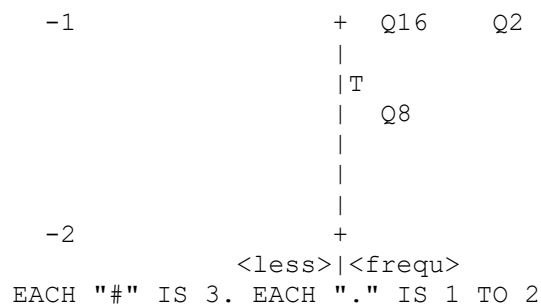
DIF was explored across gender and the results revealed that items 4, 9, and 18 show statistically significant DIF ($p < 0.05$), since their DIF contrast (the difference between the item measures in the two classifications) is less than 0.5 logits, they are negligible. (Bond & Fox, 2015). So, it can be inferred that items of the Intercultural Sensitivity Scale function in the same way across gender, and there is no biased item against male or female respondents.

Item-person Map

The distribution of items and persons are illustrated in the map in Figure 1. One property of the Rasch model is that item difficulty estimates and person ability estimates are expressed on the same metric, and they are comparable. The distribution of items and their difficulty levels are on the right side and person ability levels on the left side of the map. As we move towards the top of the map, item endorsability increases. In a similar vein, persons on top of the map have higher levels of intercultural sensitivity.

Figure 1
Wright Map of the Distribution of Persons and Items





This map provides evidence for the representativeness of the items, that is, content validity. There is no ceiling effect, because there are item thresholds whose difficulty levels are above the most able participants' ability level. Since there are gaps at the higher end of the scale, some more items are needed at that difficulty level to cover the gaps and provide a more precise estimate of the persons.

Unidimensionality Analysis

Invariance of Item Calibration across Subsamples

The main assumption of the Rasch model is unidimensionality. To check unidimensionality and generalizability aspect of construct validity, the invariance of item measures across different populations or the invariance of person measures across different sets of items can be investigated. The generalizability aspect indicates to what extent we can depend on the test scores as broader indications of individual's ability and not just as his/her ability to perform a limited number of tasks of a specific test. Messick (1989) states that test scores are considered as a function of test items, the persons who respond to the items, and the context of measurement. It is important to prove whether test scores are generalizable across contexts or not.

The generalizability aspect of construct validity was investigated by checking the invariance of item calibration across one possible set of subsamples. Control lines indicate which items do not display invariance within the boundaries of measurement error across the two- person samples. (Wright & Masters, 1982).

Figure 2 shows the bi-variate plot of item measures calibrated on the basis of two groups of high measures and low measures participants. As the plot indicates all the items are located on the line and it can be inferred that item calibration is invariant across the samples.

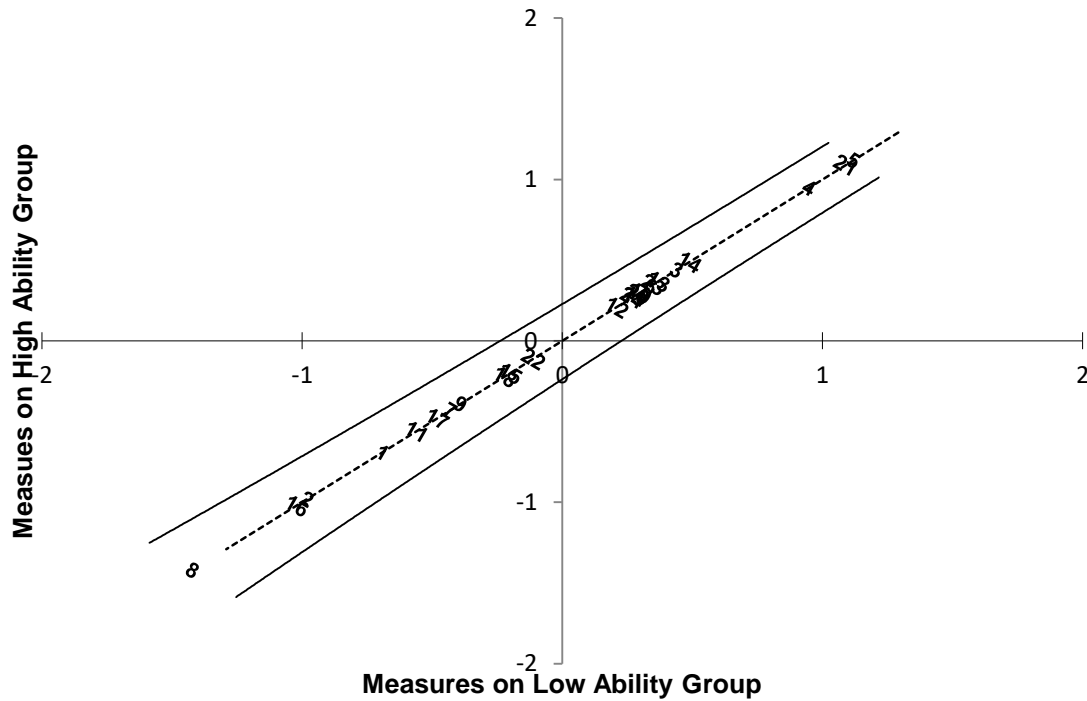
Principal Component Analysis of Standardized Residuals

Principal Component Analysis (PCA) of standardized residuals was applied to look for patterns in the data which are not in agreement with the Rasch measures or unexpected part of the data. Eigenvalues above two indicate that the dimension extracted from residuals is a secondary dimension that threatens the unidimensionality of the scale (Linacre, 2009).

The results showed that there are two contrasts with eigenvalues above two, meaning that two clusters of items function differently from the cluster measuring the intended construct. According to Linacre (2009), examination of the content of contrasting clusters of items with high positive and negative loadings (above 0.4) can guide us to discover the secondary dimensions.

Figure 2

Bivariate Plot of Item Measures for Intercultural Sensitivity Scale for Low And High Performing Group



Investigating the results indicated that there are 13 items with positive and 11 items with negative loadings. Among positively loading items, *items (9, 22, 7, 18)* and among negative loading items, *items (14, 21, and 13)* have loading values above 0.4. Deleting positively and negatively loading items separately did not result in a unidimensional scale. Therefore, it should be examined whether these two clusters are secondary dimensions or not. This can be investigated by checking the disattenuated correlations between the person parameters based on items with positive loadings and those with negative loadings. Correlations above 0.7 mean that item clusters measure the same constructs and correlations under 0.3 are considered as a poor correlation and it is indicating that item clusters measure different constructs. Here the disattenuated correlation between positively loading and negatively loading items was (0.75), which shows that clusters are measuring the same thing, but in slightly different ways.

Analyzing the content of items indicated that most of positively loading items included those items which have negative content such as “I get upset easily when interacting with people from different cultures”, “I don’t like to be with people from different cultures”, “I think people from other cultures are narrow-minded”, and conversely, negatively loading items encompassed the items which have positive content such as “I am pretty sure of myself in interacting with people from different cultures”, “I am open-minded to people from different cultures”, “I have a feeling of enjoyment towards differences between my culturally-distinct counterpart and me”. It seems that there is a secondary dimension in the intercultural sensitivity scale, but considering the wording of the items revealed that this just might be regarded a statistical artifact. Several possible mechanisms have been suggested for artifactual factors, including “lack of ability to understand negatively worded items” (Cordery & Sevastos, 1993) and “carelessness in reading items” (Schmitt

& Stults, 1985). In addition, it is noted that negatively worded items are often less reliable and valid than positively worded items probably due to the greater difficulty for people in interpreting negatively worded items. (Schriesheim et al., 1991; Schriesheim & Hill, 1981). However, it is recommended to include some negatively worded items as a check against response bias, social conformity, or dishonesty (Bond & Fox, 2007).

Discussion

This study examined Chen and Starosta's Model of Intercultural Sensitivity which was developed in the USA context. An earlier attempt to reproduce the model in Germany has been successful. The present study reached the same result based on German and US-American sample. The purpose of the current study was to validate the Persian translation of the ISS. For validating this scale, the Rasch rating scale model was utilized. The unidimensionality of the Persian translation of Intercultural Sensitivity Scale was investigated through this model. The overall results indicated that there was neither any misfitting item nor any gender DIF or biased item against male or female participants. All the category thresholds were ordered which revealed that the respondents could distinguish well between categories of the Intercultural Sensitivity Scale.

The index of separation for both items and persons showed the acceptable value, and this confirmed the external validity or generalizability of the scale. Person separation index indicated that the instrument is sensitive enough to distinguish between approximately two groups of high and low culturally sensitive participants, and the item separation index denoted that the items of Intercultural Sensitivity Scale are representative enough in distinguishing the levels of intercultural sensitivity (high, moderate, low). Also, the invariance of item calibration across subsamples was examined by dividing the sample into two groups of high and low ability persons and creating its scatter plot. The results showed that all the items were located exactly on the line and this confirmed the generalizability aspect of the construct validity and also unidimensionality. The Item-person map indicated that the scale is on-target, too, and no ceiling effect was reported for the scale. The results of principal component analysis (PCA) of standardized residuals revealed that there were two contrasts with eigenvalues above two. Deleting positively and negatively loading items separately did not change the magnitude of contrasts significantly. Thus, the disattenuated correlation between two groups of positively and negatively loading items was estimated. Since the reported value was above 0.7 it was concluded that items of the scale measure the same thing, but in various ways. Content analysis of the items revealed that most of the positively loading items included those items which had negative content and conversely, negatively loading items enclosed those items which had positive content. Hence, there are two artifact psychometric dimensions in the scale. Furthermore, the wording of the items was scrutinized and it revealed that the existence of negatively worded items in the scale created statistical artifacts and the secondary dimension here can be interpreted to be artifacts of wording.

Factors produced by items written in opposite directions have been noted for several scales. Such factors have been found in Hackman and Oldham's (1975) "Job Diagnostic Survey" (Idaszak & Drasgow, 1987), Meyer and Allen's (1984) "Affective and Continuance Commitment Scales" (Magazine et al., in press), Rizzo et al. (1970) "Role Ambiguity and Conflict Scales" (McGee et al., 1989), Baghaei et al. (2014) "Foreign Language Reading Anxiety Scale (Saito et al., 1999) , and Spector's (1988) "Work Locus of Control Scale" (Spector, 1992). In all cases these factors were regarded to be artifacts of wording direction.

Spector et al. (1997) concluded that:

Although the item direction factors in scales could be produced by independent constructs, it seems more likely that they are artifacts. An examination of the literature on the psychometric properties of item responses suggests that the way in which people respond to items that vary in direction can produce item direction factors even when the underlying construct is unidimensional. (p. 661)

Whereas it is suggested that oppositely worded items can produce artifactual factor structures, the use of such items should not be avoided. It should be noted that the use of extreme items in the scales of unidimensional constructs is often necessary despite the creation of two apparently independent dimensions. It seems that the dimensions found in psychological scales are probably produced artifactually by the way subjects respond to items, rather than by the existence of multiple constructs.

Conclusion

The aim of this study was to further validate the Intercultural Sensitivity Scale (ISS) within the Iranian cultural context. To achieve this end, validation was done through Rasch Rating Scale model and findings indicated that the Persian translation of the Intercultural Sensitivity Scale is valid for measuring the intercultural sensitivity dimensions. The results showed that Rasch Rating Scale model worked well for establishing the validity of ISS and it was demonstrated that the instrument is a unidimensional and valid culture-free scale for measuring all intercultural sensitivity dimensions. However, some limitations should be noted. One limitation of this study is that the number of participants in this study is limited. Further studies could be conducted in order to increase the diversity among graduate and undergraduate university population. Another limitation of the current study is the distribution of male and female students across the sample. In the present study, the number of male respondents to the Intercultural Sensitivity scale is about one-third of the number of females and this may bring some difficulties in interpreting gender DIF findings. Therefore, further research is needed to use more balanced samples.

Funding Details

The author received no financial support for the research, authorship, and/or publication of this article.

Disclosure Statement

The author declared no potential conflict of interest with respect to the research authorship and/or publication of this article.

References

- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*, 561-573. <http://dx.doi.org/10.1007/BF02293814>
- Baghaei, P. (2008). The Rasch model as a construct validation tool. *Rasch Measurement Transactions*, *22*, 1145-1146.
- Baghaei, P. & Cassidy, J. (2014). Validation of the Persian translation of the cognitive test anxiety scale. *SAGE Open*, *4*, 1-11. <http://dx.doi.org/10.1177/2158244014555113>

- Baghaei, P., & Shoahosseini, R. (2019). A note on the Rasch model and the instrument-based account of validity. *Rasch Measurement Transactions*, 32, 1705-1708.
- Baghaei, P., Hohensinn, C., & Kubinger, K. D. (2014). The Persian adaptation of the foreign language reading anxiety scale: A psychometric analysis. *Psychological Reports*, 114, 315-325. <https://dx.doi.org/10.2466/08.11.PRO.114k16w1>
- Bhawuk, D. P. S., & Brislin, R. (1992). The measurement of intercultural sensitivity using the concepts of individualism and collectivism. *International Journal of Intercultural Relations*, 16, 413-436. [http://dx.doi.org/10.1016/0147-1767\(92\)90031-O](http://dx.doi.org/10.1016/0147-1767(92)90031-O)
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. Lawrence Erlbaum.
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human Sciences* (3rd ed.). L. Erlbaum. <http://dx.doi.org/10.4324/9781410600127>
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, 111(4), 1061-1071. <http://dx.doi.org/10.1037/0033-295X.111.4.1061>
- Carothers, D. (2018). A culture of equality?. *Journal of Culture and Values in Education*, 1(2), 42-57. <https://doi.org/10.46303/jcve.01.02.3>
- Chen, G. M., & Starosta, W. J. (1996). Intercultural communication competence: A synthesis. *Communication Yearbook*, 19, 353-383. <http://dx.doi.org/10.1080/23808985.1996.11678935>
- Chen, G. M., & Starosta, W. J. (2000). The development and validation of the intercultural communication sensitivity scale. *Human Communication*, 3, 1-15.
- Chi, R., & Suthers, D. (2015). Assessing intercultural communication competence as a relational construct using social network analysis. *International Journal of Intercultural Relations*. <https://doi.org/10.1016/j.ijintrel.2015.03.011>
- Cordery, J. L., & Sevastos, P. P. (1993). Responses to the original and revised Job Diagnostic Survey: Is education a factor in responses to negatively worded items? *Journal of Applied Psychology*, 78, 141-143. <http://dx.doi.org/10.1037/0021-9010.78.1.141>
- Delante, N. (2020). Classroom teaching as an intercultural communication phenomenon: A thematic analysis of a foreign teacher's rhetorical practices as communication dynamic. *Journal of Intercultural Communication Research*. <https://doi.org/10.1080/17475759.2020.1790406>
- Dong, Q., Day, K., & Collaco, C. (2008). *Overcoming ethnocentrism through developing intercultural sensitivity and multiculturalism* [Paper presentation]. International Communication Association Conference, Montreal.
- Flores, H., I. A., & Palacios, M, N. (2018). Cultural and intercultural education: Experiences of ethnoeducational teachers in Colombia. *Australian Journal of Teacher Education*, 43(7). <http://dx.doi.org/10.14221/ajte.2018v43n7.4>
- Fritz, W., Graf, A., Hentze, J., Mollenberg, A., & Chen, G. M. (2005). An examination of Chen and Starosta's model of intercultural sensitivity in Germany and the United States. *Intercultural Communication Studies*, 14(1), 53-65.
- Fritz, W., Mollenberg, A., & Chen, G. M. (2001). Measuring intercultural sensitivity in different cultural context. *Intercultural Communication Studies*, 11(2). http://digitalcommons.uri.edu/cgi/viewcontent.cgi?article=1019&context=com_facpubs
- Hackman, J. R. & Oldham, G. R. (1975). Development of the Job Diagnostic Survey. *Journal of Applied psychology*, 60, 159-170. <http://dx.doi.org/10.1037/h0076546>

- Halpern, C. (2018). Multicultural for who? A review of a comprehensive multicultural textbook. *Journal of Culture and Values in Education*, 1(2), 58-62. <https://doi.org/10.46303/jcve.01.02.4>
- Idaszak, J. R., & Drasgow, F. (1987). A revision of the Job Diagnostic Survey: Elimination of a measurement artifact. *Journal of Applied Psychology*, 72, 69-74. <http://dx.doi.org/10.1037/0021-9010.72.1.69>
- Linacre, J. M. (1999). Investigating rating scale category utility. *Journal of Outcome Measurement*, 3, 103-122.
- Linacre, J. M. (2009). *A user's guide to WINSTEPS*. MESA Press.
- Linacre, J.M. (2006). *WINSTEPS Rasch Measurement Computer Program*. Chicago: Winsteps.com.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Lawrence Erlbaum Associates. <http://dx.doi.org/10.4324/9780203056615>
- Magazine, S. L., Williams, L. J. & Williams, M. L. (in press). A confirmatory factor analysis examination of reverse coding effects in Meyer and Allen's. Affective and Continuance Commitment Scale. *Educational &*
- McGee, G. W., Ferguson, C. E. Jr., & Seers, A. (1989). Role conflict and role ambiguity: Do the scales measure these two constructs? *Journal of Applied Psychology*, 74, 815-818. <http://dx.doi.org/10.1037/0021-9010.74.5.815>
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-104). Macmillan.
- Meyer, J. P., & Allen, N. J. (1984). Testing the "side-bet theory" of organizational commitment: Some methodological considerations. *Journal of Applied Psychology*, 69, 372-378. <http://dx.doi.org/10.1037/0021-9010.69.3.372>
- Moradi, E., & Ghabanchi, Z. (2019). Intercultural sensitivity: A comparative study among business English undergraduate learners in two countries of Iran and China. *Journal of Ethnic and Cultural Studies*.6(3), 134-146. <http://dx.doi.org/10.29333/ejecs/278>
- Peng, S. Y. (2006). A comparative perspective of intercultural sensitivity between college students and multinational employees in China. *Multicultural Perspectives*, 8(3), 38-45. http://dx.doi.org/10.1207/s15327892mcp0803_7
- Peng, S. Y., Rangsipah, S., & Thaipakee, S. (2005). Measuring intercultural sensitivity: A comparative study of ethnic Chinese and Thai nationals. *Journal of Intercultural Communication Research*, 34(2), 119-137.
- Rizzo, J. R., House, R. J., & Lirtzman, S. I. (1970). Role conflict and ambiguity in complex organizations. *Administrative Science Quarterly*, 15, 150-163. <http://dx.doi.org/10.2307/2391486S>
- Samovar, L. A., & Porter, R. E. (1991). *Communications between cultures*. Wadsworth.
- Satio, Y., Horwitz, E. K., & Garza, T. J. (1999). Foreign language reading anxiety. *Modern Language Journal*, 83, 202-218.
- Schmitt, N., & Stults, D. M. (1985). Factors defined by negatively keyed items: The results of careless respondents? *Applied Psychological Measurement*, 9, 367-373. <http://dx.doi.org/10.1177/014662168500900405>
- Schriesheim, C. A., & Hill, K. D. (1981). Controlling acquiescence response bias by item reversals: The effect on questionnaire validity. *Educational and Psychological Measurement*, 41, 1101-1114. <http://dx.doi.org/10.1177/001316448104100420>
- Schriesheim, C. A., Elsenback, R. J., & Hill, K. D. (1991). The effect of negation and polar opposite item reversals on questionnaire reliability and validity: An experimental investigation.

- Educational and Psychological Measurement*, 51, 67-78.
<http://dx.doi.org/10.1177/0013164491511005>
- Spector, P. E. (1988). Development of the Work Locus of Control Scale. *Journal of Occupational Psychology*, 61, 335-340. <http://dx.doi.org/10.1111/j.2044-8325.1988.tb00470.x>
- Spector, P. E. (1992). Summated rating scale construction: An introduction. Sage.
- Spector, P. E., Van Katwyk, P. T., Brannick, M. T., & Chen, P. T. (1997). When two factors don't reflect two constructs: How item characteristics can produce artifactual factors. *Journal of Management*, 23, 659-677. <http://dx.doi.org/10.1177/014920639702300503>
- Tétreault, S., Bétrisey, C., Brisset, C., Gulfi, A., Schaer, M., Leanza, Y., & Kühne, N. (2020). Intercultural Experiences Prior to the Educational Program: Occupational Therapy and Social Work Students. *Journal of Culture and Values in Education*. <https://doi.org/10.46303/jcve.2020.6>
- Vu, T. (2019). Theoretical Constructs and Practical Strategies for Intercultural Communication. *Journal of Curriculum Studies Research*, 1(1), 43-53. <https://doi.org/10.46303/jcsr.01.01.4>
- West, L. (2009). *Evaluating the intercultural sensitivity scale with counselors in international schools* [Unpublished doctoral dissertation]. Regent University.
- Wolf, E. W., & Smith, E. V. (2007). Instrument development tools and activities for measure validation using Rasch models: Part I-instrument development tools. *Journal of Applied Measurement*, 8, 204-234.
- Woofter, S. (2019). Book Review: Building Equity: Policies and Practices to Empower All Learners. *American Journal of Qualitative Research*, 3(1), 136-139. <https://doi.org/10.29333/ajqr/5815>
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. MESA Press.

Notes On Contributors

Elahe Moradi is PhD candidate at English Department, Ferdowsi University of Mashhad, Iran.