Construct Validity and Reliability of the Tolerance Scale among Iranian College Students

Ercümden Ersanlı
Department of Social Works,
Faculty of Health Sciences, Ondokuz Mayıs University, Samsun, Turkey

Shiva Saeighi Mameghani
Department of Guidance and Psychological Counseling,
Faculty of Education, Ondokuz Mayıs University, Samsun, Turkey

Abstract

In the present study, the Tolerance Scale developed by Ersanlı (2014) was adapted to the Iranian culture, and its validity and reliability were investigated in the case of Iranian college students. The participants consisted of 552 Iranian college students (62% male, $M=20.84$, $S.D.: 1.53$) selected using the convenience sampling method. The sample was randomly divided into two groups, and the item analyses and exploratory factor analysis were conducted on the data obtained from the first sample and the confirmatory factor analysis on the data from the second sample. The item analyses revealed that some items in the scale had low item-total correlations. After the removal of these items, the exploratory factor analysis demonstrated that the scale had a single-factor structure similar to that of the original scale. The confirmatory factor analyses confirmed the one factor structure on a new sample. These initial findings showed that some cultural factors might be influential in the measurement of tolerance levels. However, the single-factor structure of the scale, like that of the original, suggested that the concept of tolerance could be discussed as a psychological phenomenon with a single dimension. In further research, the factor structure of the scale can be examined on various samples, including adolescents or adults. Moreover, the predictive validity of the scale can also be investigated to test its ability to predict undesirable behaviors, such as violence.

Keywords: Tolerance Scale, validity, reliability, Iranian college students.

1. Introduction

While cultures used to have their own different, independent, and separate realms of existence, they have become more interconnected over the last century, particularly the past 25 years, due to modern and advanced means of communication and interaction (Bekiroğlu ve Şahin, 2013). Huntington (2003) states that in a world of different civilizations, each civilization has to learn to co-exist with the others because the increasing number of means of communication, more common desire to travel and to study abroad, and necessity to live in different cultures due to economic conditions make it important to empathize with people of other cultures (as cited in Bekiroğlu ve Balcı, 2014). But getting to know different cultures also entails tolerating differences.

The notion of tolerance first emerged as the concept of recognizing the rights of people and groups with different beliefs to live and of avoiding interventions in order to end religious wars in Europe. Although “tolerance” is believed to have religious connotations, it has been observed to exist in different strands of the social milieu such as ethnicity, gender, and social/class roles (Yürüşen, 1995). Groups or individuals who were considered different were often oppressed and as a result, became isolated or even worse, they were killed (Bekiroğlu ve Balcı, 2014). The peaceful co-existence of those with differences can only be achieved through mutual and highly developed tolerance (Yazgan, 2007).

“Tolerance” is defined as the conscious preference of the powerful not to use their power of terminating others because of communal differences (Ersanlı, 2014). Tolerance is a term commonly used in all spheres of life. It refers to the ability to endure anything, be it physical, chemical, or psychosocial. It can include the ability to endure stress, encumbrance, pain, and/or pressure without being harmed. Tolerance is a concept used to denote endurance and bearing in the face internal tension and stress by employing one’s own means (Ersanlı, 2014).

Tolerance is an inherent power in humans that can bring them closer together. It allows humans to seek or create common ground. It can also be defined as a “magical power” that draws people together. What underlies this definition is the fact that humans are social beings. Considering socialization is indispensable, they look for similarities so that they can come together. This search increases one’s tolerance. Thus, it can be asserted that tolerance is a sine qua non of togetherness.
Prejudices arise from unknowns. If one does not employ the tolerance mechanism to get to know a person, this may lead to misconceptions and fear. Tolerance can involve many different kinds of relationships. One group might need to learn to tolerate the other, or it may even be on an individual level. Intolerance and over tolerance may also result in the disruption of societal or interpersonal relationships, while near perfect tolerance is the ideal level. Its achievement allows individuals and societies to maintain good and sustainable mental health.

Researchers note that tolerance plays a significant part in interpersonal relationships (Ersanlı, 2013). Studies on tolerance report that a high level of tolerance offers numerous benefits for individuals (Ersanlı, 2014; Yazgan, 2007). For example, Yazgan's (2007) study on Turkish college students revealed that more tolerant students tended to display less aggressive behavior than those who were less tolerant and it was easier for them to control their anger. According to Ersanlı (2014), because tolerance helps people respect differences and act more tolerantly in different conditions, it helps people adopt a more democratic stance in interpersonal relationships. Despite the affirmative effects of tolerance on humans' lives, there are only a few measurement tools in the psychological literature to assess tolerance levels, including the Tolerance Scale developed by Ersanlı (2014).

The Tolerance Scale (TS) is a 5-point Likert scale with 11 items. The construct validity of the scale was tested by Ersanlı (2014) by conducting exploratory and confirmatory factor analysis, and the findings substantiated its single-factor structure. More precisely, Ersanlı (2014), as a result of the exploratory factor analysis, found that the scale contained only one factor, which could account for 39.37% of the total variance, and that its item factor loadings ranged between .46 and .75. The confirmatory factor analyses conducted on a different sampling substantiated its single-factor structure. The reliability data demonstrated that item-total correlations of the scale ranged between .37 and .64, and that it had an adequate level for the Cronbach’s Alpha internal consistency coefficient (α=.79). However, whether the scale would prove valid in other countries, such as Iran, had to be investigated because Turkish culture has become more individualistic due to rapid Westernization, although it has preserved its collectivistic characteristics (Mocan-Aydin, 2000). But Iran has a collectivistic culture and is governed by Islamic codes (Dastmalchian, Javidan, & Alam, 2001). Although it is asserted that 98% of the population of Turkey is Muslim, it is a secular state. For example, Islamic codes dictate that a Muslim woman is forbidden to marry a Christian man who has not converted to Islam. This action is less tolerated by Iranian culture, but more by Turkish culture, although it is a Muslim country. In addition, tolerable behaviors in Turkey may not be accepted by Iranians. The present study intended to adapt the Tolerance Scale to the Iranian culture and to investigate the validity and reliability of the adapted version.

2. Method

2.1. Participants

The sample of this cross-sectional study consisted of 552 Iranian college students selected using convenience sampling. Participants included 344 (62%) male and 208 (38%) female students, ranging in age between 18 and 23 years old. The mean age was 20.84 years (S.D.: 1.53). The participants were students of universities in either Tehran or Tabriz studying in 19 different fields. Of these students, 15.8% (n=87), 9.2% (n=51), 9.2% (n=51), 7.6% (n=42), 6.7% (n=37) and 6.5% (n=36) were studying counseling and guidance, electrical and electronics engineering, civil engineering, medicine, theology, and science teaching, respectively. While 18% (n=100) were married, 82% (n=452) were single. The sample was randomly divided into two equal groups. While the item analyses and exploratory factor analysis of the scale were carried out on the first group, confirmatory factor analysis was performed on the second to determine whether the scale had a similar factor structure.

2.2. Measures

Tolerance Scale: As previously mentioned, the Tolerance Scale is a 5-point scale with 11 items. The responses to each item range from 1 (Strongly Disagree) to 5 (Strongly Agree). Ten items were reverse-scored. The total score varies between 11 and 55. Higher scores indicate higher tolerance levels. An example item from the Tolerance Scale is “I can marry a member of another culture.”

Translation of the Tolerance Scale

The translation of the scale into Persian was performed in five steps, as suggested by Brislin et al. (1973): initial translation, review of the initial translation, back-translation, review of the back-translation, and expert review. In the first step, the Tolerance Scale, originally produced in Turkish, was translated into Persian by four academicians with a proficient command of both Turkish and Persian. Then, the translations were compared, the
most acceptable version was determined, and thus the Persian form was created. In the subsequent step, two professional linguists were commissioned to translate the accepted Persian version into Turkish. Minor differences in the translations were eliminated to create the Turkish form. The back-translated Persian form was determined to be very similar to the Turkish one. In the final step, the resultant Persian form was delivered to five Persian experts in counseling and guidance to evaluate its intelligibility, and it was concluded that the scale contained no unintelligible item.

**Personal Information Form:** This form was used to obtain information about participants’ age, gender, major, and marital status.

### 2.3. Procedure

The data were collected in the fall term of the 2015–2016 academic year. Before the data collection, the required permission and approval were obtained from the relevant institutions and the ethics board. Prior to the administration of the scale, the students were assured that participation was voluntary, responses would be kept confidential, and they could discontinue the session at any time without fear of being penalized or sanctioned. The Demographic Information Form and the Tolerance Scale were administered by the researcher during the courses in groups of 25 to 50 participants. No student refused to participate in the research. They were offered no incentive. Administrations took around 20 minutes.

### 2.4. Statistical Analysis

The statistical analyses were conducted on SPSS 23 and Mplus 7.4 (Muthén & Muthén, 1998–2014). The sampling was randomly grouped into two. The division was carried out by the select random cases function of SPSS 23. The item analyses and exploratory factor analysis of the translated scale were performed on the first sample. In the second group, whether the factor structure obtained by the exploratory factor analysis was similar in a new sample was examined. The item discrimination of the scale was investigated through item-total correlations, and the items with .20 and lower item-total correlations (Nunnally & Bernstein, 1994) were removed from the data set. The Cronbach’s Alpha internal consistency coefficient was used to calculate the internal consistency of the scale. Furthermore, descriptive statistics were employed to obtain item means and standard deviations. In order to find out how many factors the scale contained, an exploratory factor analysis was conducted on the first sampling group. Prior to the exploratory factor analysis, a series of tests were performed in line with the experts’ views (Costello & Osborne, 2005; Fabrigar, Wegener, MacCallum, & Strahan, 1999; Izquierdo, Olea, & Abad, 2014) to determine whether the correlation matrix of the data was suitable for factor analysis and how many factors this correlation matrix could contain. Bartlett’s Test of Sphericity was employed to question whether there was an adequate correlation between the items. Moreover, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was utilized to find out whether the sample was adequate to run the factor analysis. Bartlett’s Test of Sphericity tests whether the correlation matrix of the data is an identity matrix. To be able to carry out factor analysis, the results of this test must be significant and the coefficient of the Kaiser-Meyer-Olkin measure of sampling adequacy must be .60 or higher (Kaiser, 1974). The number of factors with eigenvalue higher than one and scree plot were used to determine how many factors to retain. Cohesion of the items and interpretability of the obtained factors were other criteria (Hair et al., 2014). While an eigenvalue greater than unity suggests that factors with eigenvalue higher than one significantly contribute to the explained variance rate of factors, scree plot is intended for the enumeration of factors in the scale through the visual examination of eigenvalues. After the enumeration, correlation of the scale items was analyzed by the unweighted least squares factoring method. To discover which factor each item belongs to, the criterion that factor loadings of the scale items should be .30 or higher (Field, 2013) and should yield a distance of .10 or higher from the other factors was adopted.

After the factor structure of the scale was determined, the confirmatory factor analysis was carried out on the second group to find out whether the scale would yield a similar factor structure in a new sample. The mean- and variance-adjusted weighted least squares (WLSMV) estimation method was employed in the confirmatory factor analysis. The rationale behind the use of this analysis is that the WLSMV outperforms the Maximum Likelihood method in the case of ordinal Likert scales with a limited number of response categories, such as the one adapted in this study (Bandalos, 2014; Finney ve DiStefano, 2006, 2013). The suitability of the data for the model was investigated by a series of goodness-of-fit statistics in the confirmatory factor analyses (Hu & Bentler, 1999; Marsh, Balla, & McDonald, 1988): Corrected Chi-Square ($\chi^2$/df), RMSEA 90% confidence interval, CFI, Tucker-Lewis Index, Weighted Root Mean Square Residual (WRMR). For $\chi^2$/df, values below 5 and above 3 refer to an acceptable fit and those below 3 to a perfect fit. For the RMSEA, values between .08 and .06 refer to a good fit and those below .06 to a perfect fit. Moreover, for the CFI and TLI, values over .90
and .95 signify an acceptable and perfect fit, respectively. Lastly, for the SRMR, values equal to or below 1 signify a good fit (Schreiber, Nora, Stage, Barlow, & King, 2006).

3. Results

3.1. Reliability Analyses

Item and reliability analyses were carried out on the first group to test the reliability of the scale. Six items (1, 3, 5, 7, 8, 10) with an item-total correlation of .20 or lower were removed from the scale (Nunnally & Bernstein, 1994). Table 1 presents the values pertaining to the remaining items’ mean scores and standard deviations, item-total correlations, and Cronbach’s Alpha values, if an item has been removed.

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>r</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS2</td>
<td>1.82</td>
<td>.78</td>
<td>.33</td>
<td>.63</td>
</tr>
<tr>
<td>TS4</td>
<td>2.08</td>
<td>.92</td>
<td>.43</td>
<td>.59</td>
</tr>
<tr>
<td>TS6</td>
<td>1.96</td>
<td>.99</td>
<td>.46</td>
<td>.57</td>
</tr>
<tr>
<td>TS9</td>
<td>2.48</td>
<td>1.33</td>
<td>.47</td>
<td>.59</td>
</tr>
<tr>
<td>TS11</td>
<td>1.46</td>
<td>.62</td>
<td>.42</td>
<td>.61</td>
</tr>
</tbody>
</table>

Note. $r=$ Item-Total Correlation, $\alpha =$ if item deleted, N=276.

As presented in Table 1, all of the item means are less than 2.5, which marks the scale midpoint, and item standard deviations range between .62 (TS11) and 1.33 (TS9). The item-total correlations of the remaining items vary between .33 and .47. The Cronbach’s Alpha internal consistency coefficient calculated for the entire scale is .65.

3.2. Exploratory Factor Analysis

In order to find out how many factors the remaining TS items contain, an exploratory factor analysis was conducted on the first sampling group. In the literature, exploratory and confirmatory factor analysis are used to test the construct validity of scales. However, while exploratory factor analysis is a multivariate statistical method adopted to create a theoretical structure, in other words to develop a theory, confirmatory factor analysis is a multivariate statistic employed for the evaluation of an existing theory (Keith, 2005). Likewise in this study, because six TS items were removed and how many factors they would have contained was unknown, firstly an exploratory factor analysis was conducted. It was observed that Bartlett’s Test of Sphericity produced significant results ($\chi^2_{(10)}=190.52, p<.001$) and the Kaiser-Meyer-Olkin measure of sampling adequacy yielded a coefficient of .71. Upon seeing that the data were suitable for factor analysis, the correlation matrix was examined by exploratory factor analysis by using the unweighted least squares method, a factoring technique. It was discovered that the correlation matrix of the scale contained only one factor with an eigenvalue higher than one. The factor had an eigenvalue of 2.15 and was able to account for 43.04% of the total variance. It was also observed that the scree plot produced a distinct break in the single factor. Table 2 contains the item factor loadings, eigenvalues, and communality values obtained by the exploratory factor analysis.

<table>
<thead>
<tr>
<th></th>
<th>Factor Loadings</th>
<th>Communalities (Initial)</th>
<th>Communalities (Extraction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS2</td>
<td>.40</td>
<td>.16</td>
<td>.16</td>
</tr>
<tr>
<td>TS4</td>
<td>.53</td>
<td>.21</td>
<td>.28</td>
</tr>
<tr>
<td>TS6</td>
<td>.62</td>
<td>.27</td>
<td>.38</td>
</tr>
<tr>
<td>TS9</td>
<td>.60</td>
<td>.22</td>
<td>.36</td>
</tr>
<tr>
<td>TS11</td>
<td>.54</td>
<td>.21</td>
<td>.29</td>
</tr>
<tr>
<td>Eigenvalue</td>
<td>2.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explained variance (%)</td>
<td>43.04</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.3. Confirmatory Factor Analysis

A confirmatory factor analysis was performed on the second sampling group to determine whether the single-factor structure obtained by the exploratory factor analysis would have a similar structure in the case of a different sampling. In the confirmatory factor analysis, fitness of the single-factor model to the data set was evaluated. The factor variance was fixed to one for the description of the model. The analysis indicated that the single-factor model in question perfectly fit the data set ($\chi^2$ (df =5)=6.746, $p >.05$; RMSEA=.036 90% CI [.000-.096], $p >.05$; CFI=.994; TLI=.987; WRMR=.467). Table 3 presents the standardized item factor loadings, t-values, and $R^2$ values obtained by the confirmatory factor analysis.

Table 3. Results of the Confirmatory Factor Analysis

<table>
<thead>
<tr>
<th>Factor Loadings</th>
<th>t</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS2  .293 (.065)</td>
<td>4.54</td>
<td>.09</td>
</tr>
<tr>
<td>TS4  .469 (.060)</td>
<td>7.79</td>
<td>.22</td>
</tr>
<tr>
<td>TS6  .869 (.060)</td>
<td>14.51</td>
<td>.76</td>
</tr>
<tr>
<td>TS9  .511 (.058)</td>
<td>8.82</td>
<td>.26</td>
</tr>
<tr>
<td>TS11 .677 (.066)</td>
<td>10.22</td>
<td>.46</td>
</tr>
</tbody>
</table>

Note. Numbers in parenthesis are standard errors of factor loadings. All t-values were significant at $p<.001$.

As indicated in Table 3, the item factor loadings based on the confirmatory factor analysis ranged between .29 and .87, and all the t-values were significant. These results substantiate the single-factor structure obtained by the exploratory factor analysis in the case of a new sample group.

4. Discussion

In the present study, the Tolerance Scale developed by Ersanlı (2014) was adapted to the Iranian culture, and the validity and reliability of the adapted version were tested in the case of Iranian college students. Because the item-total correlations of some items were found out to be .20 or less by reliability analyses, they were removed from the scale. The fact that these items are not discriminative enough in the case of Iranian college students signifies that some cultural factors may be influential in the assessment of tolerance. For example, one of the removed items reads “I can marry a member of another culture.” Even if a great portion of the Turkish population is Muslim, the state has a secular regime. Iran is an Islamic state and governed by Islamic codes, which stipulate that for a woman to be able to marry a non-Muslim man, he has to convert to Islam. A similar situation holds true for men. Turkish society, however, has experienced dramatic changes in its values as a result of westernization and cultural changes in recent decades (Mocan-Aydin, 2000). Therefore, Turkish people have become more apt to adopt a more tolerant attitude towards such marriages than those in Iran. This may make the items less discriminative. Yet, it was observed that the remaining items had a sufficient level of discriminativeness (Nunnally & Bernstein, 1994). Moreover, Cronbach’s Alpha internal consistency coefficient of the scale was calculated to be .65. Considering that this coefficient is a value dependent on the number of items in a scale and the TS contains a limited number of items, it can be postulated that the scale has an acceptable level of reliability.

The exploratory factor analysis demonstrated that the scale had a single-factor structure and was able to account for 43% of the total variance. This finding is similar to the variance rates reported for the 11-item scale by Ersanlı (2014). It is recommended in the related literature that single-factor scales have a variance rate of 30% or higher (Büyüköztürk, 2010). Hence, it is possible to conclude that the scale is suitable for research purposes. Further, the confirmatory factor analysis showed that the factor structure of the scale was similar in a different sampling. The fact that tolerance has been proven to be a single-dimension phenomenon suggests that it may have similar structures in different cultures, but that some cultural factors should be taken into consideration when evaluating this structure.

The present study has some limitations. It investigated only the construct validity of the scale in the Iranian culture and produced partly substantiated results. However, more evidence is needed in relation to different types of validity. Thus, further research can test its convergent and divergent validity and its validity to predict certain undesirable behaviors (e.g., violence). Secondly, the construct validity of this scale was investigated only for Iranian college students. The way members of different groups perceive and interpret “tolerance” may vary. Therefore, its construct validity can be investigated on adolescents and adults in future research. Lastly,
researchers noted limitations regarding to use of self-report scales such as social desirability and mid-point responding (Şahin, Barut, & Ersanlı, 2013a, 2013b; Şahin, Barut, Ersanlı, & Kumucağız, 2014; Şahin, Ersanlı, Kumucağız, Barut, & Ak, 2014). Thus, statistical inferences regarding to factor structure of this scale may also investigate in further studies using other informant data such as close friend and family members.

The present study investigated the validity and reliability of the Tolerance Scale for the case of the Iranian culture and ascertained that it could be used in this culture. But the non-operationality of some items in its translated version, which existed in the original scale, makes it difficult to compare the results. A possible solution to this issue could be to investigate the validity and reliability of the items, which are valid and reliable in the Iranian case, in other cultures and to compare the obtained results. Another solution would be to calculate values between 1 and 5 by adding up the points for each answered item and then dividing the resultant number by the number of items, and to make the comparison based on the obtained figures rather than the total scores. Lastly, the comparison of effect sizes would indicate how significant the results are in practice.

References


