I tested the dimensionality, reliability, and convergent validity of the Marital Offence-Specific Forgiveness Scale (MOFS) with a sample of 158 married Turkish couples. The MOFS consists of 10 items used to elicit 2 dimensions, namely, resentment–avoidance and benevolence, and is designed to measure forgiveness regarding any offense, wrongdoing, or sin committed in a marital context. To test the factorial structure of the scale, confirmatory factor analysis was used. The psychometric features of the MOFS support its applicability to research conducted within the Turkish cultural context.

Keywords: Marital Offence-Specific Forgiveness Scale, marital transgression, marital offense, forgiveness, scale adaptation, reliability, validity, Turkey.

Forgiveness is one of the central concepts in philosophy, religion, psychology, ethics, and political science (Meninger, 1996), and there has been a noticeable increase in the number of forgiveness-related studies published in the field of psychology, specifically, over the last two decades (González Martín, Rodríguez González, & Génova Fuster, 2011). Forgiveness is defined as goodwill toward an offender following the initial feelings of resentment, desire for revenge, and anger (McCullough, Worthington, & Rachal, 1997). Although there is not one common definition of forgiveness in the literature, researchers (e.g., Freedman, 1998; McCullough, Pargament, & Thoresen, 2000) are quite clear that forgiveness is not synonymous with excusing or condoning a transgression, and that it is possible,
and sometimes advisable, for forgiveness to occur without reconciliation. In the related literature, scholars discuss both the benefits and costs of forgiveness in the context of marriage. As regards the benefits of forgiveness, many researchers (Fincham, Beach, & Davila, 2004; Gordon, Baucom, & Snyder, 2005; McNulty, 2008) discuss the notions of well-being within relationships, marital longevity, and marital satisfaction. On the other hand, in a growing body of research (e.g., McNulty, 2008; Zimbardo, 1970) the main focus has been on the costs of forgiveness for both partners in the relationship. In these studies, the researchers have discussed the role of the context in which the frequency of the partner’s transgressions is the main factor. The more frequent the transgressions, the lower is the level of accountability that is necessary to motivate behavior change.

Discussions of appropriate methods and data collection procedures are also very common in the current literature. For example, a study conducted by Paleari, Regalia, and Fincham (2005) was criticized by McNulty (2008) because Paleari et al. used a cross-sectional method in their study and, thus, the findings could not reflect the long-term effects of forgiveness. In terms of scales dealing with the subject of forgiveness that are in current use, international studies have led to the popularizing of a number of instruments to assess this concept. Two examples mentioned by Paleari et al. (2005) are the Marital Forgiveness Scale (Fincham, Beach, & Davila, 2004) and the Forgiveness Inventory (Gordon, 2003), which are favored because they are multi-item scales and, thus, reflect more dimensions of offense-specific marital forgiveness than do single-item self-reported measures (Afifi, Falato, & Weiner, 2001; Karremans, Van Lange, Ouwerkerk, & Kluwer, 2003). Still other instruments used to assess forgiveness are the Transgression-Related Motivation Inventory (TRMI; McCullough et al., 1998) and the Forgiveness Scale (Rye et al., 2001).

Forgiveness positively affects physical and mental health (Exline & Baumeister, 2000; Huang & Enright, 2000), while the chronic state of unforgiveness is associated with negative health outcomes (Berry & Worthington, 2001; Wade & Worthington, 2003). Therefore, long-term forgiveness behaviors must be explored by counselors, mental health practitioners, and academics. More specifically, forgiveness has not been well studied in Turkey, with the exception of three recent studies (i.e., Alpay, 2009; Ergüner-Tekinalp & Terzi, 2012; Taysi, 2007), indicating that there is a gap in the Turkish cultural context that this study can start to fill.

Recently, Taysi (2007) adapted the TRMI for use in a Turkish cultural context. It has been noted that this scale can only be used to measure forgiveness based on a single offense in a mutual relationship. However, the results obtained do not provide information about whether or not the person subjected to the offense treats the perpetrator of the offense in a benevolent manner (Paleari et al., 2009; Taysi, 2007). Although the scale was developed in order to meet a need, scholars
still claim it is inadequate because it is more suitable for assessing painful situations, such as infidelity, that lead to the termination of relationships, rather than daily issues, like resentment or unfair criticism (Paleari et al., 2009).

Thus, no psychometrically robust measure of forgiveness in relationships had been developed and, to address this, Paleari, Regalia, and Fincham (2009) presented psychometric data for the development of the Marital Offence-Specific Forgiveness Scale (MOFS). My aim in this study was to adapt and examine the psychometric properties of a Turkish version of the MOFS.

**Method**

**Participants**

In this study, I gathered data from 170 married couples; however, because some participants did not answer all items, data from only 158 couples were included in the analysis. These couples (comprising 158 women and 158 men) were located in the cities of Mersin, Adana, and Antakya in Turkey, and were selected via a snowball sampling technique. Specifically, volunteer participants were parents and relatives of students in the Faculty of Education at Mersin University. The average age of the married women was 38.32 years ($SD = 9.52$) and of the married men was 41.13 years ($SD = 10.37$). The duration of the marriages ranged from 3 months to 41 years, with an average marriage duration of 14.5 years. Participants were all in their first marriage and had a minimum of one child.

**Measures**

**The Marital Offence-Specific Forgiveness Scale (MOFS).** Paleari et al. (2009) developed the MOFS, in which responses to the items are measured on a 7-point Likert scale ($1 = $strongly disagree$ to $7 = $strongly agree$). Of the original 12 items, the benevolence dimension comprises five, the resentment dimension comprises five, and the avoidance dimension comprises two. The original scale was tested as one-, two-, and three-factorial versions and, after explanatory and confirmatory factor analyses were conducted, the 10-item, two-factorial version was chosen. In the two-factorial design, items 1, 3, 4, 6, 7, and 8 elicit resentment–avoidance (e.g., “Since my wife/husband behaved in that way, I have been less willing to talk to her/him” and “Since my wife/husband behaved in that way, I get annoyed with her/him more easily”) and items 2, 5, 9, and 10 (e.g., “Although she/he hurt me, I definitely put what happened aside so that we could resume” and “Since my wife/husband behaved that way, I have done my best to restore my relationship with her/him”) elicit benevolence, with these items retaining the same meaning for both men and women. The new form of the scale has acceptable internal consistency, with Cronbach’s alpha values for each dimension being as follows: resentment–avoidance = .80 (men) and .75 (women), and benevolence = .80 (men) and .75 (women).
Beck Depression Inventory (BDI). The BDI was developed by Beck, Ward, Mendelson, Mock, and Erbaugh (1961), and in 1981, Hisli undertook validity and reliability studies in Turkey, revealing that the cut-off point of 17 and above discriminated 90% of depression. The BDI consists of 21 items that are used to assess depressive behavior, based on a 4-point Likert-type self-evaluation statement with anchors of 0 = *not at all like me* and 3 = *exactly like me*.

Rosenberg Self-Esteem Scale (RSES). The RSES was developed by Rosenberg in 1965 and comprises 63 items classified into 12 subcategories that are used to measure self-esteem. Çuhadaroğlu (1986) completed validity and reliability studies in Turkey and the validity coefficient was determined to be $r = .71$. The test–retest method was used to establish the reliability and resulted in a coefficient of $r = .75$. For the purpose of this study, I used only the first 10 items, which elicit self-respect.

**Procedure**

The English version of the MOFS was initially translated into Turkish by four academics employed within the Mersin University English language teaching department, who had spent at least 1 year in an English-speaking country, and who had an advanced level of English as a foreign language. The scale was then retranslated into English by two other instructors working in the same department. The translations that most closely reflected the original items in the MOFS were included in the Turkish version of the scale. Thus, the language validity of the final versions of the items and the scale were established.

The participants were each given the MOFS, BDI, and RSES in an envelope and asked to fill in the MOFS based on the latest offense by their spouse to which they had been exposed. Individual participants filled in the scale apart and independently from their spouse. After the completion of each scale, the sealed envelopes were given back to the researcher. After 1 month, the participants again completed the MOFS as a retest, and they were asked to remember the offense that they had mentioned in the first test. I had instructed participants to use the same name or nickname on each copy of the test and retest forms, and the forms of each participant were matched for statistical analysis by this means.

**Data Analysis**

To test the factorial structure of the MOFS, a confirmatory factor analysis (CFA) was run using LISREL version 8.51. The Pearson product-moment correlation coefficient was used for item–total test correlation calculations, and
to assess the test–retest reliability. Cronbach’s alpha reliability coefficient was used to assess internal consistency. SPSS version 17.0 was used for all analyses except the structural validity studies.

Results

Structural and Criterion-Related Validity Assessments
To test whether or not the structural pattern created by Paleari et al. (2009) was valid for the Turkish sample, I ran a CFA. The significance levels of the model fit coefficients, path diagram, and factor loadings were investigated separately for the groups of men and women, using the test and retest data. Thus, the structural pattern was controlled for both data sets in terms of time and gender. First, a CFA was performed on 12 items, as in the original MOFS, to assess the appropriateness of the eliminated items to a Turkish cultural context; however, the goodness-of-fit indices were below the acceptable level for both the one- and two-factor solutions. Next, a one-factor solution with 10 items was used. The goodness-of-fit indices resulting from the CFA are shown in Table 1.

Table 1. Ten-Item One-Factorial Confirmatory Factor Analysis of Compatibility Coefficients

<table>
<thead>
<tr>
<th>Group</th>
<th>Application</th>
<th>$\chi^2(df), p$</th>
<th>RMSEA [CI]</th>
<th>NFI</th>
<th>CFI</th>
<th>GFI</th>
<th>AGFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>Test</td>
<td>171.33(35), .001</td>
<td>.16 [.130, .180]</td>
<td>.56</td>
<td>.60</td>
<td>.82</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td>Retest</td>
<td>98.33(35), .001</td>
<td>.12 [.093, .140]</td>
<td>.76</td>
<td>.83</td>
<td>.88</td>
<td>.81</td>
</tr>
<tr>
<td>Women</td>
<td>Test</td>
<td>108.00(35), .001</td>
<td>.13 [.100, .150]</td>
<td>.69</td>
<td>.76</td>
<td>.86</td>
<td>.79</td>
</tr>
<tr>
<td></td>
<td>Retest</td>
<td>85.07(35), .001</td>
<td>.10 [.076, .130]</td>
<td>.72</td>
<td>.81</td>
<td>.90</td>
<td>.84</td>
</tr>
</tbody>
</table>

Note. $\chi^2(df) =$ chi square (degrees of freedom), RMSEA = root mean square error of approximation, NFI = normed fit index, CFI = comparative fit index, GFI = goodness-of-fit index, AGFI = adjusted goodness-of-fit index.

Results set out in Table 1 show that the model setup was not confirmed, that is, that the one-factorial model did not have acceptable data compatibility. To test whether or not the scale was compatible with the suggested two-factorial model, a CFA was performed on 10 items. The goodness-of-fit statistics are presented in Table 2, and illustrate that the fit of the two-factorial structure was confirmed for both men and women. To check the measurement equivalence of the two-factor model for husbands and wives, another CFA was conducted for both the test and retest applications. The fit statistics for test measurement equivalence were as...
Figure 1. Test measurement equivalence path diagrams of the 10-item, two-factorial model.
Figure 2. Retest measurement equivalence path diagrams of the 10-item, two-factorial model.
follows: $\chi^2 = 134.65$, $df = 68$, $p < .001$, RMSEA = .078, CI = [.060, .100], NFI = .81, CFI = .90, GFI = .92. Path diagrams of this model are illustrated in Figure 1. The fit statistics for retest measurement equivalence were as follows: $\chi^2 = 105.50$, $df = 68$, $p < .001$, RMSEA = .06, CI = [.030, .080], NFI = .85, CFI = .94, and GFI = .93. Path diagrams of this model are given in Figure 2. Oblique factors were allowed for all two-factorial solutions. Satisfactory fit indices were obtained for measurement equivalence at both pretest and posttest. For all two-factorial oblique solutions, negative correlation coefficient estimates, ranging from -0.66 to -0.38, were obtained for the latent variable relationships. The item analysis and reliability calculations based on the confirmed model are shown below.

<table>
<thead>
<tr>
<th>Group</th>
<th>Application</th>
<th>$\chi^2 (df, p)$</th>
<th>RMSEA [CI]</th>
<th>NFI</th>
<th>CFI</th>
<th>GFI</th>
<th>AGFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>Test</td>
<td>68.73(34), .001</td>
<td>.079</td>
<td>.82</td>
<td>.90</td>
<td>.92</td>
<td>.87</td>
</tr>
<tr>
<td></td>
<td>Retest</td>
<td>51.60(34), .027</td>
<td>.063</td>
<td>.88</td>
<td>.95</td>
<td>.93</td>
<td>.89</td>
</tr>
<tr>
<td>Female</td>
<td>Test</td>
<td>65.92(34), .001</td>
<td>.081</td>
<td>.81</td>
<td>.90</td>
<td>.92</td>
<td>.87</td>
</tr>
<tr>
<td></td>
<td>Retest</td>
<td>53.89(34), .016</td>
<td>.057</td>
<td>.82</td>
<td>.92</td>
<td>.94</td>
<td>.90</td>
</tr>
</tbody>
</table>

### Item Analysis and Reliability Assessments

Results set out in Table 3 show the item–total test correlations, Cronbach’s alpha internal consistency coefficient, and the stability index of the test–retest reliability coefficient for the 10 items and two dimensions of the scale. It is clearly demonstrated that, with the exception of item 6, the values related to the item–total test correlations were above .30 and, thus, had an acceptable level of discriminatory validity. Even though the values of item 6 were below .30, it was still within the .20–.29 range; as such, I chose not to eliminate item 6 from the scale. Verification of this item could be undertaken in future studies.

Cronbach’s alpha coefficient values for the resentment–avoidance factor ranged between .67 and .73. Generally, values above .70 are considered to be at an acceptable level. Cronbach’s alpha coefficient values for the benevolence factor of the scale ranged from .66 to .73. Although the values changed based on the groups, the benevolence factor was considered to be consistent because the majority of values were above .70. Although no test–retest analysis was performed on the original scale, in this study, the test–retest coefficients for both factors were found to range from .66 to .72. As a result, both factors can be said to lend themselves to stable measurement.
### Table 3. *Item Reliability Analysis*

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Item no.</th>
<th>Men</th>
<th></th>
<th></th>
<th></th>
<th>Women</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ITTC</td>
<td>α</td>
<td>ITTC</td>
<td>α</td>
<td>ITTC</td>
<td>α</td>
<td>ITTC</td>
<td>α</td>
</tr>
<tr>
<td>Resentment–avoidance</td>
<td>1</td>
<td>.432</td>
<td>.488</td>
<td>.403</td>
<td>.432</td>
<td>.403</td>
<td>.432</td>
<td>.403</td>
<td>.432</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.594</td>
<td>.560</td>
<td>.532</td>
<td>.488</td>
<td>.532</td>
<td>.488</td>
<td>.532</td>
<td>.488</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.479</td>
<td>.375</td>
<td>.317</td>
<td>.412</td>
<td>.317</td>
<td>.412</td>
<td>.317</td>
<td>.412</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>.537</td>
<td>.594</td>
<td>.546</td>
<td>.442</td>
<td>.546</td>
<td>.442</td>
<td>.546</td>
<td>.442</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>.395</td>
<td>.545</td>
<td>.528</td>
<td>.384</td>
<td>.528</td>
<td>.384</td>
<td>.528</td>
<td>.384</td>
</tr>
<tr>
<td>Benevolence</td>
<td>2</td>
<td>.282</td>
<td>.441</td>
<td>.397</td>
<td>.459</td>
<td>.397</td>
<td>.459</td>
<td>.397</td>
<td>.459</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>.381</td>
<td>.669</td>
<td>.384</td>
<td>.454</td>
<td>.384</td>
<td>.454</td>
<td>.384</td>
<td>.454</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>.581</td>
<td>.567</td>
<td>.496</td>
<td>.555</td>
<td>.496</td>
<td>.555</td>
<td>.496</td>
<td>.555</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>.589</td>
<td>.550</td>
<td>.566</td>
<td>.483</td>
<td>.566</td>
<td>.483</td>
<td>.566</td>
<td>.483</td>
</tr>
</tbody>
</table>

*Note.* ITTC = item–total test correlation.
Criterion-Based Validity Assessment

Karremans, Van Lange, and Holland (2005) indicated that individuals who lack forgiveness may experience feelings of loneliness and depression, and may have some pathological tendencies, such as emotional deprivation. The aim in the original MOFS was to elicit the relationship between variables, such as loneliness and depression, which are considered to be indicators of well-being. Paleari et al. (2009) used responsibility attributions, rumination, emotional empathy, marital quality, closeness, marital support, depression, and self-esteem scales for assessing criterion-based validity, whereas in the current study, I employed the BDI and RSES, which had previously been adapted for use in a Turkish cultural context.

During the development of the original MOFS (Paleari et al., 2009), the resentment dimension correlated (.33) with the Center for Epidemiologic Studies Depression Scale (CES-D) in men. In this study, too, the resentment–avoidance dimension correlated with the BDI, with values of .20 ($p < .05$) for the test and .33 ($p < .01$) for the retest. In terms of the benevolence dimension, in the original scale, the correlation with the CES-D was -.29, and, in this study, it was -.03 ($p > .05$) for the test and -.17 ($p < .05$) for the retest.

During the development of the original MOFS, a negative relationship was found between the resentment–avoidance dimension in men and the RSES (-.25); in the current study, the values reflecting a similar relationship were -.11 ($p > .05$) for the test and -.15 ($p > .05$) for the retest. For the original scale, the relationship between the RSES and benevolence dimension was reported as .21. In the current study, the benevolence dimension and its correlation with the RSES was determined to be .07 ($p > .05$) for the test and .07 ($p > .05$) for the retest. Although there are similarities in the criterion-related validity correlation coefficients, weaker correlations were obtained for this study.

During the development of the original MOFS, a positive relationship (.35) was found between the resentment–avoidance dimension and the CES-D in women. In the current study, the values reflect a similar relationship of .31 ($p < .01$) for the test and .35 ($p < .01$) for the retest. In the original MOFS, the value of -.24 ($p > .05$) suggested a relationship between the benevolence dimension and the CES-D. In the current study, the benevolence–BDI relationship was -.04 ($p > .05$) for the test and -.03 ($p > .05$) for the retest.

During the development of the original scale, the resentment–avoidance dimension was found to correlate (-.33) with the RSES for males. In the current study, the resentment–avoidance and RSES correlations were -.006 ($p > .05$) for the test and -.064 ($p > .05$) for the retest. The benevolence dimension in the original scale showed a positive relationship (.28) with the RSES, whereas in the current study, the values for benevolence and RSES were negative, at -.049 ($p > .05$) for the test and -.090 ($p > .05$) for the retest. In general terms, even though...
the criterion-based validity correlation coefficients were similar to those arrived at during the analysis of the original scale, the coefficients related to RSES did not indicate a significant correlation between the two test results.

Discussion

In this study, I adapted the MOFS (Paleari et al., 2009) for use with a Turkish population and examined forgiveness of an offense within a marriage relationship with a group of Turkish married couples. The CFA performed to examine the structural validity of the scale revealed two valid and reliable factors, as in the original scale, namely, resentment–avoidance and benevolence. Paleari et al. (2009) mentioned that in this two-factor structure, the presence of benevolent and conciliatory motivation toward the offender cannot be inferred from the absence of resentful and avoidant motivation, and, likewise, the lack of benevolence does not imply the existence of resentful and avoidant motivation toward the partner. In testing the structural validity of the scale, the CFA results suggested a factorial structure parallel to that found in the original scale. In agreement with the current literature, it can be claimed that to measure forgiveness in a marriage context, both the nonexistence of negative reactions and the existence of positive reactions toward the transgressor need to be examined. The two-factorial nature of the structure suggests that feelings such as resentment and avoidance can coexist in marital relationships. My results in this study, which support those of previous researchers, show that married couples have the following characteristics: a) they simultaneously experience resentment toward, and demonstrate physical/psychological avoidance of, the transgressor, and b) they use avoidance as a way of indicating resentment (Fincham, Hall, & Beach, 2005, 2006; Paleari et al., 2009). In the original MOFS scale, the results of the CES-D and RSES were examined in relation to other variables, as indicators of well-being. To account for parallelism with the original scale, I also used the BDI and RSES to assess the criterion-based validity. In general terms, even though the criterion-based validity correlation coefficients were similar to those arrived at during the analysis of the original scale, my results did not indicate a correlation with the RSES.

Future researchers should expand this study by sampling from a more diverse population. In addition, a much larger sample and the assessment of different variables will help strengthen the validity of this instrument. My analysis results with regard to the MOFS indicate that the internal consistency of the Turkish adaptation shows similarity to that of the original MOFS. During the original development of the scale, a test–retest reliability analysis was not performed, so I undertook this in the current study. For both factors, the 1-month test–retest reliability showed that both factors are consistent in terms of measurement. In
terms of whether the MOFS can be used to differentiate between the reactions of men and women to a marital offense, in relation to the results in my study, it can only be said that both the men and women exhibited similar behavior patterns in terms of resentment–avoidance and benevolence. This result aligns with the research findings of McCullough et al. (1998), Taysi (2007), and Paleari et al. (2009).

This study has several limitations that restrict the generalizability of the findings. The participants were all living in cities in the south of Turkey. Future researchers should determine whether or not the MOFS performs adequately in different part of the country. A second limitation is that all participants tended to be in well-adjusted marital relationships. Future researchers should also assess couples in distressed marital relationships or divorced couples. The third limitation is the reliance on self-report measures. As McCullough, Hoyt, and Rachal (2000) pointed out, forgiveness measures that go beyond self-report are still needed. In addition, the effects of variables such as ethnicity, gender, and severity of the transgression on MOFS scores should be determined in a Turkish cultural context. Clearly there is still much work to be done with respect to the measurement of forgiveness. In spite of these limitations, my adaptation of the MOFS appears to have adequate psychometric properties and, thus, represents an important step in the development of the marital forgiveness literature in Turkey.

References


