Abstract
This study aimed to adapt the Emotional Regulation Questionnaire into Turkish and investigate its validity and reliability. The Turkish version of the Emotion Regulation Questionnaire was administered to 442 university students. Firstly, language equivalence of the Turkish version of the Emotion Regulation Questionnaire was investigated and findings indicated that the Turkish translation of the Emotion Regulation Questionnaire is equivalent to the original form. Confirmatory factor analyses successfully revealed the two-factor (reappraisal and suppression) structure of the Turkish version the Emotional Regulation Questionnaire. Further, based on the two-factor model, results from multi-group analysis indicated that the Turkish version of the Emotional Regulation Questionnaire has the property of strict factorial invariance across gender, including equal factor loadings, equal item intercepts, and equal item uniqueness (error variances and covariances). The internal consistencies were .78 for the reappraisal and .73 for the suppression subscales. Test-retest reliabilities were .74 for the reappraisal and .72 for the suppression subscales. Overall, the Turkish version of the Emotional Regulation Questionnaire has adequate psychometric properties and can be used to examine emotion regulation among Turkish university students.

Keywords: Emotion regulation; reappraisal; suppression; validity; reliability.

Introduction
Emotional regulation includes internal and external processes responsible for managing people’s emotions toward goal achievement and the manipulation of both positive and negative emotions. Process does not target only reducing the intensity and frequency of an emotional response, but also generating and preserving an emotional response. Process-oriented model provides a deeper theoretical background conceptualizing emotional regulation. Process-oriented model suggests that

---

1 Associate Professor, Uludağ University, Faculty of Education, Department of Educational Sciences, email; eldelek@uludag.edu.tr
2 Assistant Professor, Bayburt University, Bayburt Education Faculty, Department of Educational Sciences, email; yeroglu45@gmail.com
the generation of emotions happens over time, therefore emotion regulation strategies can be
categorized according to their temporal location along the emotion generative process. At general
level, strategies are classified as antecedent-focused, in which intervention occurs early and is focused on
changing the effect of emotion-generating cues, and response focused, referring to those which
involves managing emotions after they occur and are being experienced (Gross & Thompson,
2006).
Although process-oriented model includes several emotional regulation strategies, two of these,
have been differently operationalized in undergraduate students and community samples of adults.
These strategies are cognitive reappraisal, an antecedent-focused strategy which includes reinterpreting
an emotion eliciting event in order to change its emotional impact and expressive suppression, a
response-focused strategy that involves hindering the external facial, bodily, or behavioral signs of
the emotion (Gross, 2002).

Emotion Regulation Questionnaire (ERQ; Gross & John, 2003) was utilized to assess the habitual
use of two emotion regulation strategies. Previous studies in different countries (Balzarotti, John, &
Gross, 2010; Gross & John, 2003; Melka, Lancaster, Bryant, & Rodriguez, 2011) demonstrated that
the ERQ is a reliable and valid scale for undergraduate students. However, in Turkey, only one
study (Yurtseven, 2008) addressed this issue and reported satisfactory internal consistencies for
both subscales. Therefore, it can be said that more extensive studies are needed to conduct with the
aim of examining validity and reliability of the ERQ in Turkey.

The present study was undertaken to evaluate the language equivalence; the factorial and criterion-
related validities; the measurement invariance; the internal consistency; and the test-reliability of the
Turkish version of the Emotional Regulation Questionnaire (T-ERQ). We hypothesized that the
two-factor structure of the ERQ would be confirmed in Turkish undergraduate students, and we
anticipated that the cognitive reappraisal subscale of the T-ERQ would be positively correlated with
the reappraisal subscale of the 18-item Turkish version of the Cognitive Emotion Regulation
Questionnaire (18-item CERQ-TR) and the expressive suppression subscale of the T-ERQ would
be negatively correlated with focus on thought subscale of the 18-item CERQ-TR. Furthermore,
we anticipated the T-ERQ have measurement invariance across gender, and show acceptable
internal consistency [Cronbach’s α>.70] and test-retest reliability.

Method
Participants
The validity and reliability studies of the ERQ were carried out on four different research groups.
In the formation of these research groups, the convenience sampling method was used since it
provided the researcher with convenience in terms of time and access. The scale's linguistic
equivalence studies were carried out on the first group. This group was made up of 30 participants having masters or doctoral degrees in the field of psychological counseling and guidance. The mean age of the participants (13 male and 17 female) was 26.4 (SD=3.24). The second study group was made up of 442 students enrolled at various faculties of Uludag University, Turkey. 201 (45.5%) of the students were enrolled at the Faculty of Education, 112 (25.3%) at the Faculty of Veterinary Science, 116 (26.2%) at the Faculty of Economics and Administrative Sciences, 3 (0.7%) at the Faculty of Engineering, and 10 (2.3%) at the Faculty of Medicine. The mean age of this group, made up of 332 female (75.1%) and 110 male (24.9%) students, was 20.93 (SD=2.05). Of the participants, 23 (5.2%) came from lower socioeconomic status, while 392 (88.9%) came from middle socioeconomic status, and 26 (5.9%) were from higher socioeconomic status. Of the study group on whom the criterion-related validity studies were conducted, 101 (41.7%) were enrolled in the Education Faculty, 71 (29.3%) in the Faculty of Economics and Administrative Sciences, 58 (24%) in the Faculty of Veterinary Sciences, 2 (0.8%) in the Faculty of Engineering, and 10 (4.1%) in the Faculty of Medicine. The mean age of the participants – 172 (71.1%) of whom were male and 70 (28.9%) female – was 20.84 (SD=2.06). Of the participants, 13 (5.4 %) came from lower socioeconomic status, 214 (88.8%) came from middle socioeconomic status, and 14 (5.8%) were from higher socioeconomic status. The group on whom the test-retest studies were done was made up of 63 (70%) male and 27 (30%) female students with a mean age of 19.60 (SD=2.23). During the administration, the participants were informed about the aims of the study and that the data would be kept secret and that they could leave the study whenever they wanted. Moreover, each participant was asked for their informed consent. The participants were volunteers and received no academic credit for their participation in the study.

**Measures**

The ERQ was developed by Gross and John (2003). The ERQ is composed of 10 items assessing individuals’ habitual use of two emotion regulation strategies, i.e. reappraisal (6 items) and suppression (4 items). Responses are given on a 7-point Likert-type scale ranging from strongly disagree (1) to strongly agree (7). Here is an example of the items in the reappraisal sub-dimension: “When I want to feel more positive emotions, I change what I’m thinking about.” Here is an example of the items in the suppression sub-dimension: “I keep my emotions to myself.” The lowest score which can be received from the reappraisal sub-dimension is 6, and the highest is 42. The lowest score which can be received from the suppression sub-dimension is 4, and the highest is 28. However, a total score cannot be obtained from the scale. The internal reliability coefficient of the ERQ ranges between .80 and .82 for the reappraisal factor, and between .73 and .76 for the suppression factor. Test-retest reliability over an interval of three months was .69 for both tests.
(Gross & John, 2003). A study conducted by Balzarotti et al. (2010) found that the suppression sub-dimension was negatively related with the “venting of emotions” sub-dimension of the Coping Orientations to Problems Experienced scale (r=-.48, p<.01), and the reappraisal sub-dimension was positively related with the “positive reinterpretation” sub-dimension of the Coping Orientations to Problems Experienced scale (r=.45, p<.01).

The 18-item CERQ-TR was developed by Çakmak and Çevik (2010) on the basis of the CERQ (Garnefski, Kraaij, & Spinhoven, 2001), which has 36 items evaluating cognitive aspects of emotion regulation. This scale is composed of 9 sub-dimensions, each of which includes two items, namely self-blame, acceptance, focus on thought, positive refocusing, planning, reappraisal, perspective, catastrophizing, and other-blame. The possible score range on each subscale goes from 2 to 10, where a higher score on each subscale indicates more frequent use of that particular strategy. The total score of the 18-item CERQ-TR, however, cannot be calculated. Participants indicated the degree to which they agree with each item on the scale using a 5-point Likert-type scale that ranged from 1 (Never) to 5 (Always). Cronbach alpha values for the internal consistency for each subscale were .64 for self-blame, .69 for acceptance, .63 for focus on thought, .74 for positive refocusing, .68 for planning, .68 for reappraisal, .63 for perspective, .68 for catastrophizing, and .70 for other blame (Çakmak & Çevik, 2010). In this study, Cronbach alpha internal consistency coefficients were found to be .77 for focus on thought and .87 for reappraisal.

**Procedure**

Permission was received from James S. Gross to adapt the ERQ to Turkish. Translation of the ERQ from English to Turkish was carried out in several stages. First, the scale was translated into Turkish and the quality of the translation was discussed by 4 people: the researcher, a psychological counselor, a measurement and evaluation expert, and a Turkish professor of English language and literature. After they reached agreement on the best version, it was back-translated to English by a person bilingual in Turkish and English, who had been informed about the topic in detail, but had never seen the English version of the scale. This process was repeated until the back-translation matched the original ERQ. Finally Turkish and the English versions of the ERQ were administered to 30 people with a good command of English who had either a master’s degree or doctorate in the field of psychological counseling, and the relationship between both versions was evaluated using Pearson’s correlation method. The feedback received from these 30 people and the Pearson correlation coefficients affirmed that the content of the Turkish translation of the ERQ is equivalent to the original form. Then the validity and reliability testing procedure started.

Confirmatory factor analysis (CFA) was used to test the 2-factor theoretical structure. Before applying CFA, preliminary analyses were conducted to ensure there was no violation of the
assumption of multivariate normality, homoscedasticity, and multicollinearity. The error term of each indicator was assumed to be uncorrelated with each of the others. $\chi^2/df$ (chi square to the degree of freedom ratio), CFI (comparative fit index), RMSEA (root mean square error of approximation), and GFI (goodness of fit index) were used as fit indexes. As a rule of thumb, for a model with an acceptable fit to the data, these indexes should be $\leq 5$, $>.90$, $<.08$ and $>.90$, respectively, whereas the corresponding values for a good fit would be $\leq 2$, $>.95$, $<.06$ and $>.95$, respectively (Hu & Bentler, 1999).

To investigate whether the T-ERQ has measurement invariance across gender, analyses were done by hierarchically nesting the models to conduct systematic comparison test. In other words, testing for factorial invariance of confirmatory factor analysis is substantiated by testing sequence of models, beginning with an unconstrained model and progressively introducing equality constraints on parameters. The following hierarchical ordering of nested models was tested to reflect the relative importance of the different sets of parameters: (a) unconstrained model, (b) factor loadings invariant, (c) factor loadings and error variances invariant across groups, (d) factor loadings, error variance, and factor variances invariant. In order to compare relative fit of nested models, change in CFI ($\Delta$ CFI) was used here, with a recommendation that support for the more parsimonious model (the more highly constrained model with fewer estimated parameters) can fit data better than a less parsimonious model requires a change in CFI is smaller or equal to .01 (Cheung & Rensvold, 2002; Vandenberg & Lance, 2000).

Under the criterion-related validity studies, the relationships between the reappraisal and suppression sub-dimensions of the T-ERQ and the subscales of the 18-item CERQ-TR were examined.

For the reliability study, internal consistency coefficient of the scale (Cronbach alpha) was calculated. To calculate test-retest reliability, the Pearson correlation coefficient was calculated for the relationships between the administrations held with 90 people with three-week intervals. LISREL 8.54 was used for the confirmatory factor analysis and multi-group confirmatory factor analysis and SPSS 11.5 was used for the other analyses.

**Results**

**Linguistic equivalence**

Within the scope of the linguistic equivalence studies, first of all, permission was received from James S. Gross through e-mail. Translation of the ERQ from English to Turkish was carried out in various stages. These stages based on Hambleton (1994)’s suggestions. Hambleton (1994) suggested that back translation method where the scale is translated into the target language and then it is re-translated back to source language and then people who are expert on both source and target
languages are inquired to compare the original version of the scale with the adjusted version of the scale.

Based on these stages, first, the scale was translated to Turkish and the quality of the translation was discussed by 4 people: the researcher, a psychological counselor, a measurement and evaluation expert, and a Turkish professor of English language and literature. After they reached agreement on the best version, it was back-translated to English by a person bilingual in Turkish and English, who had been informed about the topic in detail, but had never seen the English version of the scale. This process was repeated until the back-translation matched the original ERQ. Then the Turkish and English versions of the scale were administered twice to each of the 30 people at an interval of three weeks, and the relationship between both versions was evaluated using Pearson’s correlation method. Correlations between the Turkish and English versions for the subscales of reappraisal and suppression were found to be .83 (p<.001) and .87 (p<.01), respectively. The obtained correlation coefficients indicated that the T-ERQ achieved linguistic equivalence.

**Reliability**

Cronbach’s alpha coefficients for the reappraisal and suppression subscales of the T-ERQ were .78 and .73, respectively. As shown in Table 1, the item-total correlations were between .47 and .61 for the reappraisal subscale, and between .44 and .64 for the suppression subscale. The test-retest reliability was .74 for the reappraisal and .72 for the suppression subscales. That all the reliability coefficients belonging to the reappraisal and suppression dimensions (Cronbach’s α) were found to exceed .70 (Nunnally & Bernstein, 1994) indicated that their internal consistencies were high. We can also state that the test-retest reliability coefficients belonging to the reappraisal and suppression met the criterion of .70 (Nunnally & Bernstein, 1994).

Table 1. Corrected item-total correlations of all items

<table>
<thead>
<tr>
<th>Item no.</th>
<th>English item ERQ and T-ERQ</th>
<th>Corrected item-total correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reappraisal Items</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>When I want to feel more positive emotions (such as joy or amusement), I change what I'm thinking about.</td>
<td>.50</td>
</tr>
<tr>
<td>3.</td>
<td>When I want to feel less negative emotion (such as sadness or anger), I change what I'm thinking about.</td>
<td>.53</td>
</tr>
<tr>
<td>5.</td>
<td>When I'm faced with a stressful situation, I make myself think about it in a way that helps me stay calm.</td>
<td>.47</td>
</tr>
<tr>
<td>7.</td>
<td>When I want to feel more positive emotion, I change the way I'm thinking about the situation.</td>
<td>.56</td>
</tr>
</tbody>
</table>
Confirmatory factor analysis

CFA was used to confirm the 2-factor theoretical structure. According to the results of this confirmatory factor analysis, the ratio of $\chi^2$/df was found to be 1.94, CFI. 98, RMSEA .046, and GFI 0.99. According to Hu and Bentler (1999), if the ratio of $\chi^2$/df is below 2, the CFI and GFI are over .95, and RMSEA is smaller than .06, this indicates perfect harmony. Accordingly, the two-factor theoretical structure of the ERQ can be said to have given perfect harmony. Moreover, all the factor loads belonging to the two factors were observed to meet the criterion of .40 (Tabachnick & Fidell, 1996). The factor loadings of all 10 items on the reappraisal and suppression subscales are presented in Figure 1.

![Factor Loadings Diagram](image_url)
Measurement invariance of the T-ERQ across gender

To investigate whether the T-ERQ are invariant across gender, CFI differences (Δ CFI) between hierarchically nested models were analyzed on the basis of a recommendation that support for the more parsimonious model (the more highly constrained model with fewer estimated parameters) fitting the data better than a less parsimonious model requires a change in CFI is smaller or equal to .01 (Cheung & Rensvold, 2002; Vandenberg & Lance, 2000). Comparing differences in CFI is accomplished by testing a sequence of models, beginning with an unconstrained model and progressively adding equality constraints on (a) factor loadings, (b) factor loadings and error variances and (c) factor loadings, error variances and factor correlations (Brown, 2006; Çetin, 2010; Şekercioğlu, 2010; Vandenberg & Lance, 2000; Wang, Willett, & Eccles, 2011).

Table 2. Goodness of fit in confirmatory factor model across gender for the T-ERQ

<table>
<thead>
<tr>
<th>MODEL</th>
<th>χ²/df</th>
<th>RMSEA (90% CI)</th>
<th>CFI</th>
<th>NNFI</th>
<th>Δ CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1*</td>
<td>1.96</td>
<td>0.078 (0.060; 0.099)</td>
<td>0.91</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>Model 2b</td>
<td>1.82</td>
<td>0.074 (0.055; 0.094)</td>
<td>0.91</td>
<td>0.92</td>
<td>0.00</td>
</tr>
<tr>
<td>Model 3c</td>
<td>1.61</td>
<td>0.064 (0.044; 0.083)</td>
<td>0.92</td>
<td>0.91</td>
<td>0.01</td>
</tr>
<tr>
<td>Model 4d</td>
<td>1.59</td>
<td>0.063 (0.043; 0.082)</td>
<td>0.92</td>
<td>0.90</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Note: * Unconstrained b Factor loadings invariant c Factor loadings and error variances invariant d Factor loadings, error variances and factor variances invariant

Table 2 summarizes the results of measurement invariance of the model across gender for the T-ERQ. First, a multi-group analysis with the unconstrained model demonstrated an acceptable baseline model for both males and females (χ²/df=1.96, RMSEA=0.078, CFI=0.92, NNFI=0.92). Later, to test the equivalence of factor loadings across gender, factor loadings were constrained to be equal across the two groups. Multi-group analysis revealed that this constrained model was acceptable (χ²/df=1.82, RMSEA=0.074, CFI=0.92, NNFI=0.92). Also the value of CFI difference test (Δ CFI=0.00) indicated that factor loadings of both gender groups were equivalent. In addition to the factor loadings, error variances of each item were also constrained to be invariant across gender. Multi-group analysis demonstrated that this constrained model was acceptable (χ²/df=1.61, RMSEA=0.064, CFI=0.92; NNFI=0.91). Furthermore, the value of CFI difference test (Δ CFI=0.01) demonstrated that error variances of each item were invariant across gender as well as factor loadings. Finally, in addition to constraints mentioned above, factor variances were also constrained to be equal across the two groups. Multi-group analysis revealed this constrained model was acceptable (χ²/df=1.59, RMSEA=0.063, CFI=0.92; NNFI=0.90). Additionally, the value of CFI difference test (Δ CFI=0.01) demonstrated that the factor loadings, error variances, and factor variances were invariant across gender.
Criterion-related validity

The reappraisal subscale of the T-ERQ correlated positively with the reappraisal ($r=.79, p<.01$), acceptance ($r=.16, p<.01$), positive refocusing ($r=.26, p<.01$), planning ($r=.29, p<.01$), and perspective ($r=.22, p<.01$) subscales of the 18-item CERQ-TR. However, the suppression subscale of the T-ERQ only correlated negatively with the “focus on thought” subscale of the 18-item CERQ-TR ($r=-.73, p<.01$), while suppression subscale of the T-ERQ were not correlated with acceptance, positive refocusing, planning, perspective, and reappraisal subscales of the 18-item CERQ-TR.

Discussion

This study aiming to adapt the ERQ (Gross & John, 2003) to Turkish employed the translation-back translation method, one of the most frequently methods used in inter-cultural questionnaire adaptations. We followed Hambleton’s (1994) suggestion to do a pilot study subsequent to the translation procedure to detect potential difficulties which may appear in the translation. This was done using a group of 30 master’s degree and doctoral students who had a sufficient level of English proficiency. In line with the feedback we received, necessary corrections were made. Moreover, that the correlation coefficients obtained as a result of administering the Turkish and English versions of the ERQ to this group were found to be higher than .70 (Tabachnick & Fidell, 1996) (.83 for reappraisal and .87 for suppression) indicates that the T-ERQ achieved linguistic equivalence.

An internal consistency above .70 is considered adequate (Nunnally & Bernstein, 1994); thus, both the reappraisal and suppression sub-dimensions of the T-ERQ are reliable. Again, the internal consistency coefficients obtained in this study were similar to those obtained in the original study (.79 for the reappraisal and .73 for the suppression) (Gross & John, 2003) and in its Italian adaptation (.84 for the reappraisal and .72 for the suppression) (Balzarotti et al., 2010). Moreover, in the study by Yurtsever (2008), Cronbach's alpha internal consistency coefficient was found to be .83 for the reappraisal and .93 for the suppression sub-dimension. The consistency over time between the first and second administrations done over three-week intervals were found as .74 for the reappraisal and .72 for the suppression subscales. The test-retest reliability coefficients of the reappraisal and suppression sub-dimensions met the criterion of .70 (Nunnally & Bernstein, 1994). Again, this showed a similarity with the original ERQ study (.69 for the reappraisal and suppression subscales) (Gross & John, 2003). In a study by Balzarotti et al. (2010) on the Italian version of the ERQ, test-retest reliability coefficients made with a two-week interval were found to be .67 for the reappraisal sub-dimension and .71 for the suppression sub-dimension. Moreover, the corrected-
item total correlations' having been found higher than .40 (Nunnally & Bernstein, 1994) indicates that the discriminant power of the items was sufficient.

The goodness-of-fit index values obtained from the CFA shows that the fit of the model was sufficient (Hu & Bentler, 1999). The results of the CFA indicated that the items loaded on both factors above .40, which confirms the adequacy of the scale (Tabachnick & Fidell, 1996). The T-ERQ was also found to have measurement invariance across gender, indicating that the power of the relationships between each item and the underlying construct is the same for males and females. This result was consistent with Melka et al.’s (2011) study. The results from Melka et al.’s (1994) study demonstrated that the original structure of the ERQ did not vary across gender in a sample of 1118 undergraduate students. A study which examined the measurement invariance of the ERQ across two European nations has confirmed measurement invariance of the ERQ across Italian and German undergraduate student samples (Sala et al., 2012). Spaapen et al. (2014) found that the original factor structure of the ERQ was not sustained by either the Australian or United Kingdom non-student samples. Nonetheless, in their study, a well model fit was achieved for both samples when one item was removed. Furthermore, the revised ERQ-9 was found to be invariant across the samples, gender, age, and education level.

The reappraisal sub-dimension of the T-ERQ correlated positively with the reappraisal subscale of the 18-item CERQ-TR which refers to thinking of attaching a positive meaning to the event in terms of personal growth. On the other hand, the suppression subscale of the T-ERQ related negatively with the “focus on thought” subscale of the 18-item CERQ-TR, which refers to thinking all the time about the feelings and thoughts associated with the negative event. This is in line with previous findings. The study by Balzorotti et al. (2010) found that the ERQ’s reappraisal subscale correlated positively with the Positive Reinterpretation subscale of Coping Orientations to Problem Experience (COPE; Carver et al., 1989). Carver et al. (1989) defined reinterpretation as “construing a stressful transaction in positive terms” involving an attempt to look for something good in what is happening and to learn from difficult experiences. The suppression subscale of the ERQ was negatively related to the “venting of emotions” subscale of the COPE; this subscale consists of being aware of one’s own emotional distress and letting it out (Carver et al., 1989).

Taken together, the results of this study indicate that the T-ERQ is a valid and reliable measurement tool. Nevertheless, the study sample’s having been composed of only university students limits the generalizability of its findings. In this context, it can be said that further studies will need to evaluate the factor structure of the ERQ in a community sample. Furthermore, doing research studies using the T-ERQ should improve the scale’s measuring power. Finally, investigating the relationships between such psychological constructs as the ERQ and emotional
intelligence, psychological well-being, and self-esteem will help increase the amount of knowledge related to the strategies of suppression and reappraisal.

References


