THE EMOTION REGULATION QUESTIONNAIRE: ERQ-9 FACTOR STRUCTURE AND MEASUREMENT INVARIANCE IN AUSTRALIAN AND CANADIAN COMMUNITY SAMPLES

SIMON M. RICE
ORYGEN, THE NATIONAL CENTRE OF EXCELLENCE IN YOUTH MENTAL HEALTH; CENTRE FOR YOUTH MENTAL HEALTH, THE UNIVERSITY OF MELBOURNE

MATT S. TREEBY
LA TROBE UNIVERSITY, MELBOURNE

ELON GERSH
ORYGEN, THE NATIONAL CENTRE OF EXCELLENCE IN YOUTH MENTAL HEALTH; CENTRE FOR YOUTH MENTAL HEALTH, THE UNIVERSITY OF MELBOURNE

JOHN S. OGRODNICZUK
DAVID KEELEY
UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER

Research into emotion regulation has accelerated since the development of the Emotion Regulation Questionnaire (ERQ). Despite wide use of the ERQ, few psychometric validity studies have been undertaken beyond student samples. The present study reports a confirmatory factor analysis (CFA) and measurement invariance study of the ERQ. Model comparisons were made with the revised nine-item version using adult community samples from Australia (N = 233, M = 37.42 years), and Canada (N = 247, M = 25.94 years). Results from both samples indicated poor model fit for the original ERQ, but good model fit and item invariance (for gender and education level) for the ERQ-9. The ERQ-9 predicted marginally greater variance in depression symptoms relative to the ERQ. Findings are consistent with recent research suggesting that the ERQ-9 may have improved clinical utility in assessing domains of emotion regulation. Further validation studies of the ERQ in non-student samples are recommended.

Key words: Emotion regulation; Validity; Confirmatory factor analysis; Factor structure; Gender.

Correspondence concerning this article should be addressed to Simon Rice, Orygen, The National Centre of Excellence in Youth Mental Health, 35 Poplar Road, Parkville Victoria 3052, Australia. Email: simon.rice@orygen.org.au

Emotion regulation refers to the process of controlling or influencing one’s affective state, ideally in ways that will increase the chance of helpful, rather than harmful outcomes (Gross, 2015). Humans use a variety of strategies to regulate emotion, and two of the most widely studied emotional regulation strategies are expressive suppression and cognitive reappraisal (Westerlund & Santtila, 2018). Findings from meta-analyses clearly show that expressive suppression is a maladaptive emotional regulation strategy which is associated with poorer mental health, including depression, whereas the tendency to regulate emotions by
engaging in cognitive reappraisal buffers individuals against psychopathology (Aldao, Nolen-Hoeksema, & Schweizer, 2010; Webb, Miles, & Sheeran, 2012). Accordingly, there has been growing interest in the assessment of emotion regulation tendencies across domains including mental health (Aldao et al., 2010), functioning within education settings (Xu, Fan, & Du, 2015), and in understanding risk-taking behaviors (Magar, Phillips, & Hosie, 2008).

Gross and John’s (2003) process model of emotion regulation focuses on antecedent-focused strategies (i.e., things done before emotion response tendencies have become fully activated), and response-focused strategies (i.e., strategies employed once an emotion is already underway). The Emotion Regulation Questionnaire (ERQ; Gross & John, 2003) is a widely used measure of these emotion regulation strategies, assessing the domains of cognitive reappraisal, an antecedent-focused strategy, and expressive suppression, a response-focused strategy (Balzarotti, John, & Gross, 2010).

While the factor structure of the ERQ has been validated in a number of exploratory and confirmatory factory analysis studies using student samples (e.g., Melka, Lancaster, Bryant, & Rodriguez, 2011), comparatively little validation work has been undertaken beyond student cohorts. A large validation study in a German community-based sample failed to support the factor structure of the original ERQ model (Wiltink et al., 2011). Similarly, a series of confirmatory factor analyses using ERQ data from independent community samples recruited in Australia and the United Kingdom reported the need for scale modification, namely, the omission of ERQ Item 3 (Spaapen, Waters, Brummer, Stopa, & Bucks, 2014). Consistent findings were observed in a cross-nation validation study (university students) which reported that ERQ Item 3 exhibited low factor loadings on the reappraisal subscale for both Italian and German respondents (Sala et al., 2012) and in a separate Italian validation. Item 3 exhibited significant error correlation with Item 1 (Balzarotti et al., 2010). While measurement invariance testing (differential item functioning) found the ERQ-9 to be stable across age, Spaapen et al. (2014) reported that males more strongly endorsed statements suggesting habitual use of expressive suppression relative to females. This gender difference has been observed in other studies (e.g., Li, Sang, & Zhang, 2015).

Given the ongoing and increasing level of interest in the construct of emotion regulation, there is a need to ensure the ERQ is psychometrically sound. In this short report we evaluate ERQ-9 model fit data and measurement invariance according to gender and education level, and examine whether the ERQ-9 predicts greater variance in self-reported depression symptoms relative to the original ERQ.

**METHOD**

**Participants**

**Australian sample.** Data were provided by a convenience sample of Australian adults, N = 233 (M = 37.42 years, SD = 14.22, median = 34, range = 18-80, males = 125, 54%), recruited as part of a larger study on sex differences in depression. Males (M = 38.92 years, SD = 14.56) in the sample were significantly older than females (M = 35.11 years, SD = 13.46), p = .041. Most participants identified as Caucasian (> 95%), were in a current romantic relationship (> 61%), and lived in a metropolitan area (> 65%). A total of 19.7% (n = 46) were studying full-time and 3.9% (n = 9) were studying part-time, 50.6% (n = 118) were working, 6.0% (n = 14) were looking for work, 19.4% (n = 45) selected “other” and there was one case (0.4%) with missing data; 29.5% (n = 69) of the sample completed secondary education or below,
33.5% (n = 78) had completed an undergraduate degree, and 25.8% (n = 60) had a postgraduate degree, and 11.2% (n = 26) had completed a trade qualification.

Canadian sample. Data were provided by a convenience sample of Canadian adults, N = 247 (M = 25.94 years, SD = 10.67, median = 22, range = 18-80, males = 66, 26.2%), recruited from a university community as part of a larger study examining mental health. There was no difference in mean age by gender. A total of 46.0% (n = 114) identified as white, 38.7% (n = 96) as Asian, and 8.1% (n = 20) as multiple ethnicities and 7.2% (n = 18) as other. Most (54.6%) were in a current romantic relationship. A total of 29.0% (n = 72) of the sample completed secondary education or below, 29.0% (n = 72) had completed an undergraduate degree, 9.3% (n = 23) had a postgraduate degree, 28.3% (n = 70) reported having some college education but no degree and 4.4% (n = 11) had a trade qualification. A total of 26.6% were not enrolled in post-secondary education. Two cases in which participants indicated their gender as “other” were excluded from the present analyses.

Measures

Emotion Regulation Questionnaire (ERQ). The ERQ (Gross & John, 2003) comprises two subscales and assesses the habitual use of two emotion regulation strategies: cognitive reappraisal and expressive suppression. Six items assess cognitive reappraisal (e.g., “When I’m faced with a stressful situation, I make myself think about it in a way that helps me stay calm”), and four items assess expressive suppression (e.g., “I control my emotions by not expressing them”). Responses are made on a 7-point scale from 1 = strongly disagree, 4 = neutral, to 7 = strongly agree. Higher ERQ scores indicate more frequent use of that strategy. The ERQ-9 omitted Item 3, from the cognitive reappraisal subscale, that is, “When I want to feel less negative emotion (such as sadness or anger), I change what I’m thinking about.”

Depression – Patient Health Questionnaire (PHQ-9). The PHQ-9 (Kroenke, Spitzer, & Williams, 2001) is a self-report rating scale of the nine symptoms of major depressive disorder, as specified in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013), for example, “Little interest or pleasure in doing things.” Respondents rate PHQ-9 items relative to the preceding two-week period, from 0 = not at all to 3 = almost every day. The PHQ-9 is a well validated and commonly used measure of depression severity, within both research and clinical practice.

Procedure

Both samples provided online data on a range of psychological constructs taking approximately 15 minutes to complete. The ERQ and PHQ-9 were the only two common measures. Ethical approval was provided by the Australian Catholic University ethics committee and the University of British Columbia (UBC) ethics committee. Australian participants were recruited via paid advertisements displayed to Australian members of the Facebook social networking site, in addition to the snowball technique. It was not possible to identity method of recruitment. Australian participants were not reimbursed. Canadian participants were recruited via notices posted around the UBC campus, and offered a modest reimbursement for participation ($30 Canadian Dollars). All participants were advised that their participation was voluntary and that withdrawal from the study could occur at any time prior to the submission of online data.
RESULTS

Descriptive statistics were analysed using SPSS 22.0. Confirmatory factor analysis (CFA) was conducted in R using the lavaan and semTools packages. Skewness and kurtosis values were observed within the normal range (i.e., ± 2.0; West, Finch, & Curran, 1995) and maximum likelihood estimation was undertaken. We used accepted cut-off values as recommended for evaluating CFA models, that is, standardized root mean square residual (SRMR) ≤ .080, comparative fit index (CFI) ≥ .950, Tucker-Lewis index (TLI) ≥ .950, root mean square error of approximation (RMSEA) ≤ .060 (Weston & Gore, 2006). Satisfactory internal consistency values were observed for the ERQ-9 subscales. Marginally higher Cronbach alpha values were observed for the five-item reappraisal subscale (i.e., ERQ-9), relative to the original six-item version (see Table 1). Significant correlations were observed between the ERQ-9 subscales and depression for both samples. There was no association observed between the two ERQ-9 subscales. Medium effect sizes were observed for gender differences in suppression for both the Australian sample (males: $M = 3.87, SD = 1.38$; females: $M = 3.20, SD = 1.17$; $t(231) = 3.96, p < .001, d = 0.52$) and the Canadian sample (males: $M = 3.94, SD = 1.26$; females: $M = 3.17, SD = 1.25$; $t(246) = 4.24, p < .001, d = 0.54$). There were no gender differences for reappraisal.

Model Fit

For both the Australian and Canadian data, fit indices of the original 10-item ERQ indicated marginal model fit (see Table 2). In both samples, the TLI value fell well under the cut-off of .950. In contrast, the ERQ-9 (with Item 3 omitted) approached acceptable fit, especially for the Canadian sample, where CFI and TLI values exceeded .950. When samples were analysed together in a combined analysis, the ERQ-9 achieved acceptable fit (SRMR = .043; CFI = .964; TLI = .950; RMSEA = .066), although the RMSEA was marginally above the suggested cut-off. Standardised coefficients for the ERQ-9 items are presented in Table 3, with consistent values observed across the Australian and Canadian samples. In both samples, Item 6 (“I control my emotions by not expressing them”) reported the lowest loading.

Measurement Invariance

Measurement invariance was conducted to establish configural, weak, strong, and strict invariance using accepted multiple-group procedures (Hirschfeld & von Bracher, 2014). When examining invariance according to gender and education, separate models for the Australian and Canadian data (as well as the combined dataset) were estimated. Decision criteria were based on established invariance testing guidelines in smaller samples < 300 cases (i.e., Chen, 2007), using the CFI change cut-off value of ≤ −.005 supplemented with a RMSEA change value of ≥ .025 (see Table 4). For the combined analysis with the larger sample, the more stringent criteria CFI change ≥ .010 supplemented with RMSEA change ≥ .015 was used. Invariance tests (see Table 4) indicated that loadings, intercepts and residuals were invariant by gender and education groups (i.e., strict invariance was established).

Predicting to Depression

Multiple regression analysis was used to test the ERQ and ERQ-9 subscales for predicting depression. Using the original ERQ, the two predictors explained 15.7% and 17.7% of variance in depression, respectively;
TABLE 1
Scale descriptive statistics and intercorrelations by sample

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>6.16 (5.83)</td>
<td>.91</td>
<td>6.82 (5.59)</td>
<td>.87</td>
<td>–</td>
<td>–.37***</td>
<td>–.36***</td>
<td>.20***</td>
</tr>
<tr>
<td>Reappraisal*</td>
<td>4.76 (1.11)</td>
<td>.82</td>
<td>4.97 (1.10)</td>
<td>.85</td>
<td>–.26***</td>
<td>–</td>
<td>.98***</td>
<td>.07</td>
</tr>
<tr>
<td>Reappraisal#</td>
<td>4.74 (1.07)</td>
<td>.81</td>
<td>4.97 (1.07)</td>
<td>.83</td>
<td>–.24***</td>
<td>.98***</td>
<td>–</td>
<td>.07</td>
</tr>
<tr>
<td>Suppression</td>
<td>3.56 (1.33)</td>
<td>.77</td>
<td>3.38 (1.30)</td>
<td>.79</td>
<td>.33***</td>
<td>–.10</td>
<td>–.10</td>
<td>–</td>
</tr>
</tbody>
</table>

Note. *= Reappraisal five-item; # = Reappraisal six-item (original); Australian correlations below diagonal; Canadian correlations above diagonal.
***p<.001. **p<.01.

TABLE 2
Summary of CFA model fit indices

<table>
<thead>
<tr>
<th>Sample</th>
<th>Model</th>
<th>df</th>
<th>χ²</th>
<th>p</th>
<th>χ²/df</th>
<th>SRMR</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
<th>LO 90</th>
<th>HI 90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australian</td>
<td>1. ERQ-10</td>
<td>34</td>
<td>104.76</td>
<td>&lt; .001</td>
<td>3.08</td>
<td>.060</td>
<td>.909</td>
<td>.880</td>
<td>.095</td>
<td>.074</td>
<td>.115</td>
</tr>
<tr>
<td></td>
<td>2. ERQ-9</td>
<td>26</td>
<td>61.69</td>
<td>&lt; .001</td>
<td>2.37</td>
<td>.049</td>
<td>.948</td>
<td>.928</td>
<td>.077</td>
<td>.052</td>
<td>.102</td>
</tr>
<tr>
<td>Canadian</td>
<td>3. ERQ-10</td>
<td>34</td>
<td>128.28</td>
<td>&lt; .001</td>
<td>3.77</td>
<td>.061</td>
<td>.907</td>
<td>.877</td>
<td>.106</td>
<td>.087</td>
<td>.126</td>
</tr>
<tr>
<td></td>
<td>4. ERQ-9</td>
<td>26</td>
<td>48.91</td>
<td>.004</td>
<td>1.88</td>
<td>.048</td>
<td>.973</td>
<td>.963</td>
<td>.060</td>
<td>.033</td>
<td>.085</td>
</tr>
<tr>
<td>Combined</td>
<td>5. ERQ-10</td>
<td>34</td>
<td>189.69</td>
<td>&lt; .001</td>
<td>5.58</td>
<td>.057</td>
<td>.911</td>
<td>.883</td>
<td>.098</td>
<td>.085</td>
<td>.112</td>
</tr>
<tr>
<td></td>
<td>6. ERQ-9</td>
<td>26</td>
<td>80.36</td>
<td>&lt; .001</td>
<td>3.07</td>
<td>.043</td>
<td>.964</td>
<td>.950</td>
<td>.066</td>
<td>.050</td>
<td>.083</td>
</tr>
</tbody>
</table>

Note. ERQ = Emotion Regulation Questionnaire; SRMR = standardized root mean square residual; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; LO = low 90% confidence interval; HI = high 90% confidence interval.
TABLE 3
Standardised item loadings for the ERQ-9

<table>
<thead>
<tr>
<th>Item (subscale)</th>
<th>Australian sample</th>
<th>Canadian sample</th>
<th>Constrained model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Loading</td>
<td>SE</td>
<td>Loading</td>
</tr>
<tr>
<td>ERQ-1 Reappraisal</td>
<td>.52***</td>
<td>.05</td>
<td>.59***</td>
</tr>
<tr>
<td>ERQ-5 Reappraisal</td>
<td>.52***</td>
<td>.05</td>
<td>.59***</td>
</tr>
<tr>
<td>ERQ-7 Reappraisal</td>
<td>.53***</td>
<td>.05</td>
<td>.56***</td>
</tr>
<tr>
<td>ERQ-8 Reappraisal</td>
<td>.84***</td>
<td>.02</td>
<td>.88***</td>
</tr>
<tr>
<td>ERQ-10 Reappraisal</td>
<td>.76***</td>
<td>.03</td>
<td>.80***</td>
</tr>
<tr>
<td>ERQ-2 Suppression</td>
<td>.76***</td>
<td>.03</td>
<td>.74***</td>
</tr>
<tr>
<td>ERQ-4 Suppression</td>
<td>.74***</td>
<td>.04</td>
<td>.87***</td>
</tr>
<tr>
<td>ERQ-6 Suppression</td>
<td>.50***</td>
<td>.06</td>
<td>.47***</td>
</tr>
<tr>
<td>ERQ-9 Suppression</td>
<td>.83***</td>
<td>.04</td>
<td>.80***</td>
</tr>
</tbody>
</table>

Note. SE = standard error.
*** p<.001.

Australia $F(2,217) = 20.07, p < .001$, reappraisal ($β = -.21, p < .001$), suppression ($β = .32, p < .001$); Canada $F(2,243) = 25.97, p < .001$, reappraisal ($β = -.38, p < .001$), suppression ($β = .22, p < .001$). When the ERQ-9 was used, the variance explained increased marginally to 16.4% and 18.2%; Australia $F(2,217) = 21.04, p < .001$, reappraisal ($β = -.23, p < .001$), suppression ($β = .31, p < .001$); Canada $F(2,243) = 26.89, p < .001$, reappraisal ($β = -.38, p < .001$), suppression ($β = .21, p < .001$).

DISCUSSION

Findings from the present analyses support and extend those of Spaapen et al. (2014), and suggest that the ERQ-9 has a more robust factor structure in a non-student sample, relative to the original 10-item version of the measure. Findings also validate the ERQ-9 in a Canadian sample comprised largely of university students. While males reported significantly higher scores for suppression, there was no mean gender difference for reappraisal in both samples. The regression models indicated that the ERQ-9 was marginally better at predicting to depression than the 10-item version. We were able to establish measurement invariance for the ERQ-9 according to gender and education level.

Regarding the omission of Item 3 from the ERQ, findings are consistent with the work of Sala and colleagues (2012) who reported low Item 3 loadings on the reappraisal factor. The negatively worded item may cause some confusion on the part of the respondent. Spaapen et al. (2014) removed Item 3 as they observed significant covariance with Item 1, with modification indices of 19.29 and 9.27 for the Australian and UK samples respectively. ERQ Item 1 emphasises the individual seeking to increasing positive affect, whereas Item 3 involves reducing negative affect. While the emotions involved are of opposite valence (i.e., joy versus sadness), the underlying reappraisal strategy is the same. Specifically, each of these items concludes with “. . . I change what I’m thinking about.” Hence, it may be that the common focus on an individual actively changing their cognition results in shared variance between these items, as opposed to the affective motivation of seeking to experience either more positive, or less negative emotion.
Recent work has suggested that desirable responding (i.e., self-deceptive enhancement) may impact responses to the ERQ (Arndt, Hoglund, & Fujiwara, 2013). We cannot rule out possible response biases from our findings (i.e., as suppression represents an unwanted characteristic, endorsing may be minimised) and this should be considered in future studies. Further, use of ecologically sensitive measures of emotion regulation and arousal (including assessment of interoceptive data) may help to understand any implicit gender differences without the potential for biased responding (Kever, Pollatos, Vermeulen, & Grynberg, 2015).
Future validity research on the ERQ should recruit larger samples from occupationally-diverse cohorts and assess the ERQ-9 against diverse constructs (including wellbeing). Such approaches should also ensure more definitive tests of model fit against established criteria, in addition to establishing the test-retest reliability, predictive validity, and external validity of the 9-item scale. Nonetheless, this study has provided further support suggesting that the psychometric properties of the ERQ-9 are superior to the ERQ-10 in both student and non-student samples. While additional validity studies are required, researchers may wish to consider use of the 9-item scale, which appears to have stronger psychometric properties than the original version. Further research and refinement of the ERQ will allow for a more robust measurement of emotion regulation tendencies.

REFERENCES


