Psychiatry Research 279 (2019) 278–283

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

Assessment of the revised Difficulties in Emotion Regulation Scales among adolescents and adults with severe mental illness

Ruby Charak[a,*], Brianna M. Byllesby[b], J. Christopher Fowler[c], Carla Sharp[d], Jon D. Elhai[e], B. Christopher Frueh[f]

[a] Department of Psychological Science, The University of Texas Rio Grande Valley, TX, United States
[b] Summa Health Traumatic Stress Center, OH, United States
[c] Houston Methodist Behavioral Health, and The University of Texas Health Science Center, Houston, TX, United States
[d] Department of Psychology, The University of Houston, TX, United States
[e] Department of Psychology, and Department of Psychiatry, University of Toledo, OH, United States
[f] Department of Psychology, University of Hawaii, HI, United States

ARTICLE INFO

Keywords:
DERS
DERS-16
DERS-SF
DERS-18
Measurement invariance
Severe mental illness
Adolescents
Adults

ABSTRACT

The Difficulties in Emotion Regulation Scale (DERS) comprising 36 items has been widely used across age, gender, psychopathology, language, and culture. Recently several alternative abridged forms have been introduced, namely, the DERS-16 (Bjureberg et al. 2016), the DERS-SF (Kaufman et al. 2016), and the DERS-18 (Victor and Klonsky, 2016), each composed of 16 or 18 items, to provide researchers and clinicians with a shorter measure of emotion dysregulation. However, no study to date has directly compared the psychometrics of these alternative forms. In the present study, using confirmatory factor analysis we first examined the factor structure of the four models of the DERS in two inpatient samples of 636 adolescents in the age-range of 12–17 years (M = 15.33, SD = 1.43), and 1807 adults in the age-range of 18–76 years (M = 34.86, SD = 14.63) with severe mental illness. Next, measurement invariance was tested comparing the two age groups across the four models of DERS. Only the DERS-SF established metric and scalar measurement invariance. Findings suggest that the factor structure of the original and the abridged models of DERS have acceptable fit, however only DERS-SF had equivalence of factor loadings and item intercepts across adolescents and adults.

1. Introduction

Emotion regulation is a transdiagnostic construct implicated in many mental disorders and contributes to the comorbidity between disorders (Kaufman et al., 2016). A frequently used measure for assessing emotion dysregulation is the Difficulties in Emotion Regulation Scale (DERS; Gratz and Roemer, 2004). This measure is a 36-item self-report questionnaire divided into six factors, namely, nonacceptance of emotional responses, difficulties engaging in goal-directed behavior, impulse control difficulties, lack of emotional awareness, limited access to emotion regulation strategies, and lack of emotional clarity (Gratz and Roemer, 2004). Because the underlying mechanism of many mental disorders involves problems regulating one's emotions, the DERS has been associated with a wide variety of psychopathology, including anxiety (Bardeen and Stevens, 2015), borderline personality disorder (Dixon-Gordon et al., 2015), depression (Dixon-Gordon et al., 2015), posttraumatic stress disorder (Lilly et al., 2014), and substance/alcohol use (Simons et al., 2017), among others. With its use across a variety of psychopathology and related phenomena, the psychometric properties of the DERS and its alternative short-forms need additional examination, especially among individuals with serious mental illness.

The DERS has been used extensively in many different populations, spanning adult and adolescent age groups. It has been used with veterans (Sippel et al., 2015), community adults (Lilly et al., 2014), college students (O'Bryan et al., 2015), outpatients (Khosravani et al., 2017), and inpatients (Fowler et al., 2014; Perez et al., 2012; Sippel et al., 2015). However, with 36 items and some over-lapping item content, it may be overly lengthy or redundant for some research questions or samples, and some studies have identified problems with its purported latent factor structure (Bardeen et al., 2012; Miguel et al., 2017). Moreover, studies have shown high correlations between the six DERS subscale scores (e.g., Perez et al., 2012). A shortened version of the measure with a consistent latent structure is needed, and recently three different abridged versions of the DERS have been proposed in the

* Correspondence to: Department of Psychological Science, The University of Texas Rio Grande Valley, Edinburg, TX 78539, United States.
E-mail address: ruby.charak@utrgv.edu (R. Charak).

https://doi.org/10.1016/j.psychres.2019.04.010
Received 12 January 2019; Received in revised form 5 April 2019; Accepted 5 April 2019
Available online 08 April 2019

literature.

First, the DERS-16 was published as a 16-item alternative version of the original DERS (Bjureberg et al., 2016). Items were selected based on item-total correlations, content validity, and reliability. It consists of five subscales, and does not contain any items from the original 'lack of emotional awareness' scale. Using a sample of outpatient women and two community adult samples, the DERS-16 demonstrated good internal reliability and test-retest reliability. Furthermore, it took significantly less time to complete, averaging 139 s, compared to 329 s for the original measure. The abbreviated items demonstrated convergent validity with clinically related behaviors (i.e., borderline personality disorder symptom severity) and emotion-related measures (i.e., experiential avoidance, etc.) relative to the full 36-item version. It may be a good alternative form of the measure; however, the authors did not examine its factor structure, and it thus needs further evaluation.

Next, the short-form version of the DERS (DERS-SF) was created, and initial validation was examined in adolescent and adult samples (Kaufman et al., 2016). This version consists of 18 items, with six factors (to be consistent with the original measure) — each with three items. The authors found that the 18-item version demonstrated similar correlation patterns relative to the full measure, and the two versions shared 81–96% of the variance. Using confirmatory factor analysis (CFA), the DERS-SF yielded good model fit, indicating good underlying latent relationships between the six factors.

Finally, an 18-item version of the DERS (DERS-18) was published (Victor and Klonsky, 2016), and this measure attempted to shorten the original DERS by half the number of items, while maintaining the original six-factor solution. Items were selected based on factor loadings from the original exploratory factor analysis (EFA) studies (Gratz and Roemer, 2004), and the highest-loading items for each of the six subscales were maintained. The DERS-18 demonstrated good internal consistency and initial validity in five samples of varying ages (adolescents and adults) and levels of psychopathology. Using EFA, the DERS-18 had adequate model fit and the six factors accounted for 78% of the overall variance. However, only a portion of their sample had a history of mental illness, so additional psychometric analyses from both adults and adolescents with serious mental illness are needed.

1.2. The current study

The DERS-16, DERS-SF, and DERS-18 have not been assessed in relation to each other. Although the DERS-SF and DERS-18 have both been tested in adolescents and adults, the DERS-16 has not been examined in adolescents. Any developmental differences across adolescence and adulthood in item responses have not been examined, which is an important gap in the psychometric properties of these measures given that the DERS has historically been used in adolescent (Perez et al., 2012) and adult samples (Gratz and Roemer, 2004). Also, the samples used so far are either not dysregulated or had mild psychopathology. Thus, the present study aimed to simultaneously compare factor structure and measurement invariance of the three short versions of DERS in adolescent and adult samples with severe mental illness. It was hypothesized that all three abridged versions of the DERS would have adequate reliability and an adequate latent factor structure. Because the shortened versions have not previously been compared against each other in adult or adolescent samples, the identification of the best fitting model was exploratory. For measurement invariance, our null hypothesis was that there would be no developmental differences (by age group) in factor loadings (metric invariance) or item intercepts (scalar invariance) on the various forms of the DERS.

2. Method

2.1. Participants

The adolescent sample comprised 636 participants in the age range of 12–17 years (M = 15.33, SD = 1.43; 64.5% girls) consecutively admitted to a specialized psychiatric hospital in the Southern U.S.A. with an average length of hospitalization of 0 to 120 (M = 35.19, SD = 14.53). Nearly 72% of participants identified as White or Caucasian, 8.9% as Hispanic or Latino, 8.9% as multiracial or other, 4.1% as Asian, 2.4% as Black or African American, and 0.8% as American Indian or Alaskan Native.

The adult sample comprised of 1807 participants in the age range of 18–76 years (M = 34.86, SD = 14.63; 48.3% female) consecutively admitted to the same specialized psychiatric hospital with an average length of hospitalization of 2–238 days (M = 44.95, SD = 21.86). Self-reported racial/ethnic composition was predominantly Caucasian (90.8%), multiracial (5.3%), African American (1.7%), Asian (1.3%), American Indian (0.3%) and Pacific Islander (0.4%), and 6.3 percent identified as being of Hispanic or Latino ethnicity. Education level was above the national average with 88.0% indicating some college experience. The majority (62.0%) of participants were not working in the 30 days prior to admission.

2.2. Procedure

Adult data were collected as part of the hospital's ongoing Adult Outcomes Project to assess treatment response. All measures used in the current study were collected within 72 h of admission to a specialty treatment unit. Assessments were conducted via hospital-wide web survey on laptop computers. This project was a hybrid clinical quality and research outcomes project; accordingly, all assessments were designed and implemented as an element of routine clinical care and integrated into treatment planning and monitoring of progress such that less than 4% of patients declined participation. Patients and their treatment teams were provided with profile scores and feedback within 24 h with the expressed intention that individual patient profiles would be used to inform treatment decisions. Patients and teams were informed that the findings would be used to evaluate the overall effectiveness of treatment and for research purposes. Similarly, adolescent data were collected as part of the adolescent outcomes project to assess treatment response; however, data were collected within two weeks of admission, with similar high rates of consent. The [Baylor College of Medicine, and the University of Houston] Institutional Review Boards approved use of the project's data.

2.3. Measures

2.3.1. The Difficulties in Emotion Regulation Scale (DERS)

The DERS comprises 36 items that load onto six subscales as mentioned previously (Gratz and Roemer, 2004; available at http://bit.ly/ders2004). See Table 1 for model item descriptions. Participants are asked to indicate how often the items apply to them, with responses ranging from 1 to 5, where 1 = *almost never*, 2 = *sometimes*, 3 = *about half the time*, 4 = *most of the time*, and 5 = *almost always*. The DERS has high internal consistent (r = 0.93), good test-retest reliability (r = 0.88, $p < .01$), and adequate construct and predictive validity (Gratz and Roemer, 2004). The internal reliability for the full scale for adolescents and adults was 0.95. The internal reliability for the subscales ranged from 0.85 to 0.92 for adolescents and adults.

2.3.2. Difficulties in Emotion Regulation Scale-16

The DERS-16 has 16 items comprising five subscales, with no items from the original lack of emotional awareness subscale (Bjureberg et al., 2016). These items have been taken from the original DERS (Gratz and Roemer, 2004; see Table 1 for the questionnaire items). The scale has high internal consistency across three samples of women from the community and in psychiatric care (α = 0.92–0.94), and good test-retest reliability (p = .85, $p < .001$), and adequate construct validity (Bjureberg et al., 2016). The internal reliability for the full scale for adolescents was 0.93 and for adults was 0.94. The internal

Table 1
Factor loadings of the four competing models of the Difficulties in emotion regulation scales among adolescents and adults.

Items	DERS-36 Adolescents/Adults	DERS-16	DERS-SF	DERS-18
Lack of awareness				
2. I pay attention to how I feel (r)	0.87/0.86		0.86/0.87	0.84/0.85
6. I am attentive to my feelings (r)	0.86/0.91			0.85/0.88
8. I care about what I am feeling (r)	0.78/0.76		0.79/0.78	
10. When I am upset, I acknowledge emotions (r)	0.72/0.70		0.71/0.69	0.66/0.65
17. When I am upset, I believe that my feelings are valid and important (r)	0.71/0.56			
34. When I'm upset, I take time to figure out what I'm really feeling (r)	0.51/0.61			
Lack of emotional clarity				
1. I am clear about my feelings (r)	0.76/0.73			
4. I have no idea how I am feeling	0.83/0.75		0.84/0.77	0.79/0.72
5. I have difficulty making sense out of my feelings	0.89/0.85	0.92/0.82	0.93/0.89	0.90/0.87
7. I know exactly how I am feeling (r)	0.79/0.77			
9. I am confused about how I am feeling	0.72/0.78	0.72/0.80	0.74/0.81	0.73/0.75
Non-acceptance of emotional responses				
11. When I'm upset, I become angry with myself for feeling that way	0.87/0.80			
12. When I'm upset, I become embarrassed for feeling that way	0.82/0.84		0.81/0.82	0.81/0.80
21. When I'm upset, I feel ashamed with myself for feeling that way	0.90/0.90	0.81/0.82		0.90/0.92
23. When I'm upset, I feel like I am weak	0.78/0.84	0.70/0.79		
25. When I'm upset, I feel guilty for feeling that way	0.84/0.84		0.84/0.82	0.82/0.83
29. When I'm upset, I become irritated with myself for feeling that way	0.91/0.89	0.88/0.85	0.93/0.91	
Impulse control difficulties under distress				
3. I experience my emotions as overwhelming and out of control.	0.89/0.86			
14. When I'm upset, I become out of control	0.89/0.85	0.89/0.83	0.89/0.88	0.86/0.81
19. When I'm upset, I feel out of control	0.90/0.89	0.89/0.84		
24. When I'm upset, I feel like I can remain in control of my behavior (r)	0.72/0.59			
27. When I'm upset, I have difficulty controlling my behaviors	0.94/0.87	0.89/0.79	0.96/0.92	0.94/0.88
32. When I'm upset, I lose control over my behavior	0.95/0.87		0.97/0.91	0.94/0.89
Limited access to emotion regulation strategies				
15. When I'm upset, I believe that I will remain that way for a long time	0.85/0.83	0.80/0.77		0.85/0.80
16. When I'm upset, I believe that I will end up feeling depressed	0.84/0.83	0.83/0.80	0.83/0.85	0.86/0.83
22. When I'm upset, I believe I can find a way to eventually feel better (r)	0.65/0.61			
28. When I'm upset, I believe that there is nothing I can do to make myself feel better	0.83/0.82	0.76/0.76	0.84/0.83	
30. When I'm upset, I start to feel very bad about myself	0.87/0.88	0.81/0.82		
31. When I'm upset, I believe that wallowing in it is all I can do	0.81/0.72			0.75/0.67
35. When I'm upset, it takes me a long time to feel better	0.81/0.81		0.83/0.82	
36. When I'm upset, my emotions feel overwhelming	0.88/0.88	0.78/0.81		
Difficulties in goal directed behavior				
13. When I'm upset, I have difficulty getting work done	0.86/0.87	0.80/81	0.87/0.88	0.82/0.82
18. When I'm upset, I have difficulty focusing on other things	0.88/0.87	0.83/0.83	0.90/0.89	0.88/0.87
20. When I'm upset, I can still get things done (r)	0.64/0.60			
26. When I'm upset, I have difficulty concentrating	0.88/0.84		0.91/0.86	0.87/0.83
33. When I'm upset, I difficulty thinking about anything else	0.92/0.89	0.83/0.82		

Note: All factor loadings were significant at $p < .05$. (r) = reverse-scored item.

reliability for the subscales ranged from 0.80 to 0.92 for adolescents, and 0.80 to 0.90 for adults.

2.3.3. Difficulties in Emotion Regulation Scale-Short Form (DERS-SF)

This abridged version of DERS consists of six subscales, each with three items, for a total of 18 items (Kaufman et al., 2016) taken from the original DERS (Gratz and Roemer, 2004; see Table 1 for the questionnaire items). The scale has high internal consistency across two samples of adolescents and college students (α = 0.89–0.91), and it had comparable concurrent validity to the original DERS (Kaufman et al., 2016). The internal reliability for the full scale for adolescents was 0.90 and for adults was 0.90. The internal reliability for the subscales ranged from 0.79 to 0.94 for adolescents, and 0.79 to −0.91 for adults.

2.3.4. Difficulties in Emotion Regulation Scale-18

This 18-item version of the DERS (Victor and Klonsky, 2016) has six subscales each consisting of three items taken from the original DERS (Gratz and Roemer, 2004; see Table 1 for the questionnaire items). The scale has high internal consistency across two samples of adolescents (in patient and high school students) and three samples of adults from the community and college samples (α = 0.87–0.92). It demonstrated strong convergent and concurrent validity by showing relationship with borderline personality disorder symptoms and original DERS scores, respectively (Victor and Klonsky, 2016). The internal reliability for the full scale for adolescents was 0.90 and for adults was 0.90. The internal reliability for the subscales ranged from 0.82 to 0.94 for adolescents, and 0.82 to −0.91 for adults.

Noteworthy is that only the DERS-36 was administered; for the three abridged versions, items were selected from the DERS-36 across both the samples, in order to conduct analyses below.

2.4. Data analyses

Missing data were treated using maximum likelihood (ML) estimation with a pairwise present approach. First, confirmatory factor analysis (CFA) was conducted using Mplus 7 software on four separate models for the DERS in order to determine the best model fit. Items were treated as ordinal because the DERS has five ordinal response options (Wirth and Edwards, 2007). Therefore, we used a polychoric covariance matrix, probit regression coefficients, and weighted least squares estimation with a mean- and variance adjusted chi-square (WLSMV). Residual error covariances were fixed to zero, and factor variances were fixed to one; all unstandardized factor loadings were freely estimated. Model fit was examined via multiple fit indices, specifically comparative fit index (CFI), Tucker-Lewis index (TLI), and root-mean-square error of approximation (RMSEA; Kline, 2016). Based on Hu and Bentler (1999) we determined excellent model fit based on CFI and TLI values greater than 0.95 and RMSEA values less than 0.06.

Table 2
Model fit information and comparison of the shortened DERS models among adults and adolescents with severe mental illness.

Model	χ^2 (*df*)	CFI/TLI	RMSEA (90% CI)	χ^2 /*ΔCFI*
DERS-16				
A	4302.66 (188)	0.754/0.686	0.134 (0.130–0.137)	A vs. B 43.99***/0.002
B	4346.65 (199)	0.752/0.701	0.131 (0.127–0.134)	A vs. C 310.37***/ 0.016*
C	4613.03 (210)	0.736/0.699	0.131 (0.128–0.134)	B vs. C 266.38***/ 0.018*
DERS-SF				
A	943.32 (240)	0.965/0.955	0.049 (0.046–0.052)	A vs. B 30.83**/0.001
B	974.15 (252)	0.964/0.956	0.048 (0.045–0.052)	A vs. C 96.07***/0.004
C	1039.39 (264)	0.961/0.955 (0.046–0.052)	0.049	B vs. C 65.24***/0.003
D	998.69 (261)	0.963/0.957	0.048 (0.045–0.051)	B vs. D 24.54**/0.001
DERS-18				
A	3925.28 (240)	0.781/0.721	0.112 (0.109–0.115)	A vs. B 0.98/0.001
B	3926.26 (252)	0.782/0.735	0.109 (0.106–0.112)	A vs. C 282.74***/ 0.015*
C	4208.02 (264)	0.766/0.729	.111 (0.108–0.114)	B vs. C 281.76***/ 0.016*

Note. Model A is a configural invariance model with no parameter constrained equal across age; Model B is a metric invariance model with factor loadings constrained to be equal; Model C is a scalar invariance model with item intercepts and factor loadings constrained to be equal. Model D is scalar invariance model with three intercepts of items freely estimated, and the remaining item intercepts and factor loading constrained to be equal.

* $p < .01$.

For the measurement invariance testing, models were computed using ML estimation with robust standard errors (MLR). The estimator was changed because WLSMV estimation does not allow for separate testing of the hierarchy of constraints, and MLR can be used for categorical data with five response options (Muthén et al., 2017). First, we tested for configural invariance (Model A) in which the factor loadings, correlations, covariances, and residual variances were allowed to vary between adolescent and adult groups. Next, factor loadings were constrained to be equal across age groups, thus testing the metric invariance model (Model B). This was followed by testing for scalar invariance (Model C) by constraining item intercepts to be equal across age groups in addition to constraining the factor loadings. Based on the suggestions by Cheung and Rensvold (2002), the change in CFI was chosen to evaluate measurement invariance. A CFI difference of <0.01 in the CFI values indicates that the models are not invariant and the null hypothesis should not be rejected. Mean differences exist when the CFI difference (ΔCFI) is greater than 0.01. Configural, metric, and scalar invariance testing was conducted between adolescents and adults (age invariance) for the three abridged versions of DERS: DERS-16, DERS-SF, and DERS-18.

3. Results

3.1. Model testing in adolescent sample

First, model testing using CFA was used to determine the model that provided the best fit for the data in the adolescent sample. *DERS-36*: The original, 36-item DERS with six factors provided acceptable fit, χ^2 (579) = 3038.64, $p < .001$, CFI = 0.94, TLI = 0.93, RMSEA = 0.08 (90% CI = 0.08–0.09). The factor loadings and correlations ranged from 0.51 to 0.95, and 0.22 to −0.79, respectively. *DERS-16*: The DERS-16, consisting of 16 items and five factors was tested. This model provided excellent fit as well χ^2 (94) = 680.13, $p < .001$, CFI = 0.97, TLI = 0.96, RMSEA = 0.10 (90% CI = 0.09–0.11). The factor loadings and correlations ranged from 0.70 to 0.92, and 0.40 to 0.80, respectively. *DERS-SF*: The Kaufman et al. DERS-SF was tested next with the 18 items specifying six factors, and this model provided excellent fit for the data χ^2 (120) = 321.94, $p < .001$, CFI = 0.99, TLI = 0.99, RMSEA = 0.05 (90% CI = 0.05–0.06). The factor loadings and correlations ranged from 0.71 to 0.97, and 0.19 to 0.74, respectively. *DERS-18:* Finally, the DERS-18 was tested with 18 items specifying six factors and demonstrated excellent fit, χ^2 (120) = 261.85, $p < .001$, CFI = 0.99, TLI = 0.99, RMSEA = 0.04 (90% CI = 0.04–0.05). The factor loadings and correlations ranged from 0.66 to 0.94, and 0.05 to 0.67, respectively. However, the three models had latent factors comprising different items and were not nested, and thus a statistical comparison of the three models is not possible.

3.2. Model testing in the adult sample

Next, the same model comparison testing was conducted using the adult sample. *DERS-36*: The original DERS-36 was found to have adequate fit, χ^2 (579) = 9995.30, $p < .001$, CFI = 0.91, TLI = 0.90, RMSEA = 0.095 (90% CI = 0.093–0.097). The factor loadings and correlations ranged from 0.71 to 0.97, and 0.19 to 0.68, respectively. *DERS-16*: It had good model fit indices, χ^2 (94) = 1750.80, $p < .001$, CFI = 0.97, TLI = 0.96, RMSEA = 0.099 (90% CI = 0.095–0.103). The factor loadings and correlations ranged from 0.77 to 0.85, and 0.51 to 0.85, respectively. *DERS-SF*: The DERS-SF was found to meet criteria for excellent model fit, χ^2 (120) = 1284.12, $p < .001$, CFI = 0.98, TLI = 0.97, RMSEA = 0.07 (90% CI = 0.07–0.08). The factor loadings and correlations ranged from 0.69 to 0.91, and 0.02 to 0.72, respectively. The factor correlation between lack of awareness and difficult in goal directed behavior was non-significant. *DERS-18*: Finally, the DERS-18 was tested. This model provided excellent model fit, χ^2 (120) = 1171.61, $p < .001$ CFI = 0.98, TLI = 0.97, RMSEA = 0.07 (90% CI = 0.066–0.073). The factor loadings and correlations ranged from 0.65 to 0.92, and 0.04 to 0.76, respectively. As mentioned previously, the three models had latent factors comprising different items and were not nested, and thus a statistical comparison of the three models is not possible.

3.3. Measurement invariance by age

Invariance testing assessed for between-group differences in metric and scalar parameter estimates of the three abbreviated models of the DERS – between adolescents ($n = 636$) and adults ($n = 1807$). Participants were coded for group membership, whereby adolescents = 1 and adults = 2. *DERS-16*: This model demonstrated metric invariance with no statistically significant difference between the metric and configural model fit indices (see Table 2). However, scalar invariance was not established (ΔCFI < 0.01; Cheung and Rensvold, 2002). This model had only equivalent factor loadings between adolescents and adults. *DERS-SF*: Measurement invariance testing demonstrated metric invariance, with no statistically significant difference between the metric and configural model fit indices (see Table 2). Also, no significant difference was found between the metric and the scalar model fit indices. Therefore, DERS-SF had scalar invariance with invariant factor loadings and item intercepts between adolescents and adults. Further, the modification indices suggested freely estimating items 12, 13, 18, and 29 (Model D in Table 2) that lead to a slight improvement in the goodness-of-fit indices (ΔCFI = 0.001). *DERS-18*: This model reached metric invariance; however scalar invariance could not be established (i.e., ΔCFI = 0.015) indicating differences between age groups on item intercepts.

4. Discussion

The aims of the present study were twofold: to examine the factor structure of the three short versions of DERS, namely, DERS-16, DERS-SF, and DERS-18, and second, to investigate if the three scales consistently had the same factor loadings and item intercepts across an adolescent and adult sample with severe mental illnesses. Confirmatory factor analyses indicated that the three abridged versions of the DERS demonstrated good fit, thus supporting the measurement model put forth by the initial investigations of the three abridged versions of the DERS (Bjureberg et al., 2016; Kaufman et al., 2016; Victor and Klonsky, 2016). Measurement invariance testing indicated that the factor structure and factor loadings of the three scales were similar across the two developmental stages of adolescence and adulthood. However, only the DERS-SF (Kaufman et al., 2016) had equivalent item mean scores (item intercepts) for adolescents and adults. In other words, while all three versions demonstrated that the observed variables (i.e., items) measure the same underlying latent constructs, and that the constructs have the same meaning across the two groups, only the DERS-SF with its scalar invariance justifies any comparison of group mean across adolescent and adults.

The confirmatory factor analyses suggested that DERS-16, DERS-SF, and DERS-18 proposed measurement model fit well in the present sample of adolescents and adults. Items loaded on their respective latent factors in all abridged versions of the DERS. The DERS-SF and DERS-18 had better goodness-of-fit indices (namely, CFI/TLI and RMSEA) than the DERS-16. The measurement invariance analyses suggested that DERS-16 and DERS-18 met metric invariance, such that, for both the scales adolescents and adults had similar factor structure and their understanding of the items comprising the latent constructs of domains of emotion dysregulation was the same. Thus irrespective of the stage of development, adolescents and adults attribute the same meaning to the underlying constructs or factors of emotion dysregulation. The DERS-SF had all the invariance properties as the other two scales, and additionally, the item intercepts were equivalent across adolescents and adults. This suggests that the mean scores across the two age groups are comparable when using DERS-SF. Findings suggest that among the shorter versions of the Difficulties in Emotion Regulation Scales, the DERS-SF can be used in studies where comparison of scores of an adolescent with the adult is required (e.g., intergenerational studies). Further, when equality constraints from two items each from the subscales of non-acceptance of emotional responses (i.e., items 12 and 29) and difficulties in goal directed behavior (i.e., items 13 and 18) were removed across adolescents and adults, the model fit improved. Further studies should carefully examine these four items and their validity against external criteria before dropping them from the DERS-SF.

Based on the results of the present study, all three of the abridged versions of the DERS demonstrated adequate model fit indices for their latent factor structure, and therefore would be appropriate to use as measures of emotion dysregulation. However, for specific empirical investigations examining emotion dysregulation across the lifespan, our results indicate the DERS-SF is best considering it is invariant across developmental span (i.e., across adolescence and adulthood). Considering previous research has used the DERS to assess emotion dysregulation domains in both adults and adolescents (Fowler et al., 2014; Perez et al., 2012) investigators should consider using a measure with consistent factor loadings and intercepts across developmental stages in order to assess emotion dysregulation most appropriately.

The present study findings should be viewed with the following limitations in mind. First, it was the 36-item DERS that was administered to adolescents and adults in the present study, and items for the shorter version of the DERS were selected from those 36 items DERS. Second, no construct validity was carried out for the abridged versions of DERS, and only pre-treatment DERS assessments were examined in the present study. Therefore, it cannot be determined if one of the shortened versions of the DERS is differentially related to psychopathology, is more or less valid relative to similar empirical constructs, or associated with clinically significant change over time. Within-person differences in DERS scores were not examined across development, but instead data were discretely divided based on age (i.e., adolescents were individuals less than 18 years old and adults were classified as all patients over 18). Measurement invariance was only examined across developmental stage, so there can be no conclusions regarding the potential differences across other demographic variables (e.g., gender). Despite these limitations the current study has several advantages. Our sample was relatively large, with a wide range of psychopathology and severe mental illness. It consisted of both adults and adolescents, allowing for comparisons based on developmental stage, which has not been examined previously in the research.

To conclude, in contrast to the 36-item DERS, the use of abridged versions can be helpful in reducing cognitive effort and respondent strain particularly in clinical practice or when conducting epidemiological research. The present findings suggest that the decision to use one abridged version over the other would depend on aims of the study. For comparison of emotion dysregulation scores among adolescents and adults, the DERS-SF would be apt as it achieved scalar invariance. Future studies should focus on examining the convergent and divergent validity of the abridged versions of DERS. Such studies should also consider the psychometric properties of the shorter DERS scales over time or their invariance across other demographic variables.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.psychres.2019.04.010.

References

Bardeen, J.R., Fergus, T.A., Orcutt, H.K., 2012. An examination of the latent structure of the Difficulties in Emotion Regulation Scale. J. Psychopathol. Behav. Assess. 34, 382–392. https://doi.org/10.1007/s10862-012-9280-y.

Bardeen, J.R., Stevens, E.N., 2015. Sex differences in the indirect effects of cognitive processes on anxiety through emotion regulation difficulties. Personal. Individ. Differ. 81, 180–187. https://doi.org/10.1016/j.paid.2014.07.009.

Bjureberg, J., Ljotsson, B., Tull, M.T., Hedman, E., Sahlin, H., Lundh, L., …, Gratz, K.L., 2016. Development and validation of a brief version of the Difficulties in Emotion Regulation Scale: the DERS-16. J. Psychopathol. Behav. Assess. 38, 284–296. https://doi.org/10.1007/s10862-01509514-x.

Cheung, G.W., Rensvold, R.B., 2002. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct. Equ. Model. 9, 233–255.

Dixon-Gordon, K.L., Weiss, N.H., Tull, M.T., DiLillo, D., Messman-Moore, T., Gratz, K.L., 2015. Characterizing emotional dysfunction in borderline personality, major depression, and their co-occurrence. Compr. Psychiatry 62, 187–203. https://doi.org/10.1016/j.comppsych.2015.07.014.

Fowler, J.C., Charak, R., Elhai, J.D., Allen, J.G., Frueh, B.C., Oldham, J.M., 2014. Construct validity and factor structure of the Difficulties in Emotion Regulation Scale among adults with severe mental illness. J. Psychiatr. Res. 58, 175–180. https://doi.org/10.1016/j.jpsychires.2014.07.029.

Gratz, K.L., Roemer, L., 2004. Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. J. Psychopathol. Behav. Assess. 26, 41–54.

Hu, L., Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct. Equ. Model. 6, 1–55. https://doi.org/10.1080/10705519909540118.

Kaufman, E.A., Xia, M., Fosco, G., Yaptangco, M., Skidmore, C.R., Crowell, S.E., 2016. The Difficulties in Emotion Regulation Scale Short Form (DERS-SF): validation and replication in adolescent and adult samples. J. Psychopathol. Behav. Assess. 38, 443–455. https://doi.org/10.1007/s10862-015-9529-3.

Khosravani, V., Bastan, F.S., Ghorbani, F., Kamali, Z., 2017. Difficulties in emotion regulation mediate negative and positive affects and craving in alcoholic patients. Addict. Behav. 71, 75–81. https://doi.org/10.1016/j.addbeh.2017.02.029.

Kline, R.B., 2016. Principles and Practice of Structural Equation Modeling, fourth ed. Guilford Press, New York, NY.

Lilly, M.M., London, M.J., Bridgett, D.J., 2014. Using SEM to examine emotion regulation and revictimization in predicting PTSD symptoms among childhood abuse survivors. Psychol. Trauma 6, 644–651. https://doi.org/10.1037/a0036460.

Muthén, L.K., Muthén, B.O., 2017. Mplus User's Guide.

Miguel, F.K., Giromini, L., Colombarolli, M.S., Zuanazzi, A.C., Zennaro, A., 2017. A brazilian investigation of the 36- and 16-item difficulties in emotion regulation scales. J. Clin. Psychol. 73, 1146–1159.

O'Bryan, E.M., McLeish, A.C., Kraemer, K.M., Fleming, J.B., 2015. Emotion regulation difficulties and posttraumatic stress disorder symptom cluster severity among trauma-exposed college students. Psychol. Trauma 7, 131–137. https://doi.org/10.1037/a0037764.

Perez, J., Venta, A., Garnaat, S., Sharp, C., 2012. The difficulties in emotion regulation scale: factor structure and association with nonsuicidal self-injury in adolescent inpatients. J. Psychopathol. Behav. Assess. 34, 393–404. https://doi.org/10.1007/s10862-012-9292-7.

Simons, R.M., Hahn, A.M., Simons, J.S., Murase, H., 2017. Emotion dysregulation and peer drinking norms uniquely predict alcohol-related problems via motives. Drug Alcohol Depend. 177, 54–56. https://doi.org/10.1016/j.drugalcdep.2017.03.019.

Sippel, L.M., Jones, R.E., Bordieri, M.J., Dixon, L.J., May, A.C., Malkin, M.L., . . ., Coffey, S.F., 2015. Interactive effects of anxiety sensitivity and difficulties in emotion regulation: an examination among individuals in residential substance use treatment with comorbid posttraumatic stress disorder. Cognit. Ther. Research 39, 245–252. https://doi.org/10.1007/s10608-014-9648-2.

Victor, S.E., Klonsky, E.D., 2016. Validation of a brief version of the Difficulties in Emotion Regulation Scale (DERS-18) in five samples. J. Psychopathol. Behav. Assess. 38, 582–589. https://doi.org/10.1007/s10862-016-9547-9.

Wirth, R.J., Edwards, M.C., 2007. Item factor analysis: current approaches and future directions. Psychol. Methods 12, 58–79. https://doi.org/10.1037/1082-989X.12.1.58.