



Which Brief Is Best? Clarifying the Use of Three Brief Versions of the Difficulties in Emotion Regulation Scale

Julie M. Skutch¹ · Shirley B. Wang² · Tom Buqo³ · Ann F. Haynos⁴ · Anthony Papa¹

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Emotion regulation is a fundamental affective process implicated in a range of clinically relevant phenomena such as mood, anxiety, and personality disorders, as well as self-harm and suicidality. Many self-report scales have been developed to measure this important construct, and the Difficulties in Emotion Regulation Questionnaire (DERS) is one of the most widely used. The DERS has extensive empirical support for its use, however, its long length impacts its utility and a briefer version is needed. Recently three brief versions of the DERS (DERS-16, DERS-SF, and DERS-18) were developed independently. Initial analyses of each of these measures found them to retain the excellent psychometric properties of the original DERS measure. However, it remains unclear which version is most ideally suited to briefly measure emotion regulation in clinical and research contexts. To clarify this point, the current study examined the existing brief DERS measures on internal reliability and concurrent validity indices in a large sample of undergraduate students ($n = 1181$). The reliability and validity of all three brief forms were found to be comparable. Additionally, if replicated, our results suggest that it may be useful for future research and clinical work to use brief versions that retain subscale scores (DERS-SF and DERS-18). Based on the results and the existing literature, we recommend that the emotion regulation field come to a consensus about which brief version to use for consistency and the ability to compare findings across studies.

Keywords Emotion regulation · Measurement · Self-report · Brief assessment

Introduction

Emotion regulation is a core affective process necessary for healthy psychological functioning (Cicchetti et al. 1995). Broadly defined, emotion regulation refers to the implicit and explicit processes used to assess and modulate emotions

in the service of goal pursuit (Thompson 1994). Research indicates that effective emotion regulation serves a protective function for those at risk of mental illness (Banyard et al. 2017; Ford et al. 2014). In contrast, emotion dysregulation, or difficulties in emotion regulation ability, is a transdiagnostic risk factor underlying the development and maintenance of a wide range of psychopathology, including mood disorders (Joormann and Siemer 2014), borderline personality disorder (BPD; Salsman and Linehan 2012), suicidal behavior (Neacsiu et al. 2017), self-harm behavior (Buckholdt et al. 2015), eating disorders (Brockmeyer et al. 2014), and substance abuse (Weiss et al. 2017).

As empirical work has elucidated the link between emotion regulation and psychological functioning, many researchers have developed measures to capture different components of emotion regulation such as cognitive emotion regulation (Garnefski et al. 2001; Gross and John 2003), interpersonal emotion regulation (Hofmann et al. 2016; Niven et al. 2011), and beliefs about regulatory ability (Hutchison and Guntherth 2013). One of the most widely used measures is the Difficulties in Emotion Regulation Scale (DERS; Gratz and Roemer 2004). The DERS was developed based upon a

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10862-019-09736-z>) contains supplementary material, which is available to authorized users.

✉ Julie M. Skutch
jskutch@unr.edu

¹ Department of Psychology, University of Nevada, Reno, Mail Stop 298, Reno, NV 89557, USA

² Department of Psychology, Harvard University, Cambridge, MA, USA

³ Department of Psychology, Hofstra University, Hempstead, NY, USA

⁴ Department of Psychiatry, University of Minnesota, Minneapolis, MN, USA

leading clinical conceptualization of emotion regulation, which emphasizes the communication and signaling function that emotions serve (Gratz and Roemer 2004; Linehan 1993). This model also prioritizes the need for flexible responses to a wide range of emotional experiences for optimum mental health. To this end, the DERS is a 36-item self-report instrument consisting of six subscales measuring difficulties in the flexible, multi-dimensional regulation of emotion, including: 1) non-acceptance of or negative reactions to emotions (non-acceptance subscale); 2) difficulties engaging in goal-oriented behavior when experiencing negative emotions (goals subscale); 3) difficulty controlling impulsive behavior when experiencing negative emotions (impulse subscale); 4) lack of emotional awareness (awareness subscale); 5) perceived inability to cope with negative emotions (strategies subscale); and 6) lack of clarity about one's emotions (clarity subscale).

DERS items were initially generated in collaboration with colleagues familiar with the emotion regulation literature (Gratz and Roemer 2004). The instrument was initially tested in two studies using undergraduate samples to examine factor structure, internal and test-retest reliability, and construct and predictive validity (Gratz and Roemer 2004). The DERS has been well-tested and there is extensive empirical support for its high internal consistency, reliability, and construct validity in undergraduate samples (Gratz and Roemer 2004; Gratz and Tull 2010), adolescents (Neumann et al. 2010), and both outpatient (Osborne et al. 2017) and inpatient (Fowler et al. 2014) clinical samples. Scores on the DERS have been associated with a range of mental health symptoms and clinically relevant behaviors (e.g., Buckholdt et al. 2015; Fowler et al. 2014; Osborne et al. 2017). Additionally, this measure has helped further research elucidating models of the development and maintenance of psychopathology (e.g., Crowell et al. 2012; Sharp et al. 2016). The DERS is also sensitive to change over time, making it ideal for use in clinical research studies that require multiple assessment points, and is frequently used in treatment outcome research (Wilks et al. 2016). However, the length of this measure may place undue burden on researchers and participants and reduce its validity, as response quality often declines with increasing duration of data collection (Galesic and Bosnjak 2009). A briefer version of the instrument with shorter administration time could increase the measure's usefulness in clinical and research settings.

For this reason, three independent efforts were recently made to shorten the original 36-item DERS (DERS-36; Gratz and Roemer 2004), all of which were found in separate analyses to retain the excellent psychometrics of the DERS-36 (Bjureberg et al. 2016; Kaufman et al. 2016; Victor and Klonsky 2016). All three scales, the DERS-16 (Bjureberg et al. 2016), DERS-SF (Kaufman et al. 2016), and DERS-18 (Victor and Klonsky 2016), were derived using different, psychometrically sound methods based largely on the initial factor analysis used in the development of the DERS-36. Two

versions (DERS-SF and DERS-18) retained subscales and one (DERS-16) did not. Construct validity for each short form was established using measures of psychopathology and clinically relevant problem behaviors believed to serve a regulatory function (i.e., self-harm, substance use, purging). Initial development of each short form is thoroughly summarized in Online Resource 1.

All three short forms have demonstrated good reliability and validity across a range of samples. However, as each has produced a slightly different brief version of the DERS (see Table 1 for item comparison), it is currently unclear which brief form should be used in clinical and research settings. The proliferation of multiple assessments to measure the same construct can lead to problems. There is evidence that even small differences between measures can lead to problems comparing effect sizes across studies (Carlson and Herdman 2012). Thus, the use of multiple brief versions of the DERS could create a scenario in which much of the literature in this area cannot be directly compared. This is particularly concerning in the case of emotion regulation research, which is already plagued by multiple discrepant definitions and assessment tools, making it difficult to collapse findings across studies (Bloch et al. 2010). Further, the availability of multiple highly similar assessments can result in confusion and inefficiencies in the research process, requiring each researcher to dedicate extra effort towards comparing the different versions of the measure before selecting one for their purposes. Thus, the use of different brief versions of the DERS is likely to foster confusion, inconsistency, and inefficiency in the emotion regulation literature.

Therefore, the current study aimed to investigate the psychometrics of the three brief versions of the DERS in the same sample to clarify the relative strength of each version in order to identify the short form best suited for future use in emotion regulation research. First, we examined the reliability of all three short forms. Second, we evaluated and compared the concurrent validity of the three brief DERS using measures of clinically relevant outcomes, including symptoms of BPD, depression, anxiety, and clinically relevant problem behaviors. Last, we explored the concurrent validity of the subscale scores for the two short forms that retained them in order to determine the utility in using a brief DERS with subscales.

Method

Participants and Procedures

Participants were undergraduate students at a university in the western United States recruited for a larger study examining emotions, the self, and relationships. In order to be eligible, participants had to be: (1) 18 years or older, (2) a native English speaker, and (3) eligible to earn research credit in a psychology classroom. Following informed consent,

Table 1 Items contained within three brief DERS

DERS-36 Items	Items Included		
	DERS-16	DERS-SF	DERS-18
Nonacceptance Subscale			
11. When I'm upset, I become angry with myself for feeling that way			
12. When I'm upset, I become embarrassed for feeling that way.		x	x
21. When I'm upset, I feel ashamed with myself for feeling that way.	x		x
23. When I'm upset, I feel like I am weak.	x		
25. When I'm upset, I feel guilty for feeling that way.		x	x
29. When I'm upset, I become irritated with myself for feeling that way.	x	x	
Goals Subscale			
13. When I'm upset, I have difficulty getting work done.	x	x	x
18. When I'm upset, I have difficulty focusing on other things.	x	x	x
*20. When I'm upset, I can still get things done.			
26. When I'm upset, I have difficulty concentrating.		x	x
33. When I'm upset, I have difficulty thinking about anything else.	x		
Impulse Subscale			
3. I experience my emotions as overwhelming and out of control.			
14. When I'm upset, I become out of control.	x	x	x
19. When I'm upset, I feel out of control.	x		
*24. When I'm upset, I feel like I can remain in control of my behaviors.			
27. When I'm upset, I have difficulty controlling my behaviors.	x	x	x
32. When I'm upset, I lose control over my behaviors.		x	x
Awareness Subscale			
*2. I pay attention to how I feel.		x	x
*6. I am attentive to my feelings.			x
*8. I care about what I am feeling.		x	
*10. When I'm upset, I acknowledge my emotions.		x	x
*17. When I'm upset, I believe that my feelings are valid and important.			
*34. When I'm upset, I take time to figure out what I'm really feeling.			
Strategies Subscale			
15. When I'm upset, I believe that I will remain that way for a long time.	x		x
16. When I'm upset, I believe that I'll end up feeling very depressed.	x	x	x
*22. When I'm upset, I know that I can find a way to eventually feel better.			
28. When I'm upset, I believe that there is nothing I can do to make myself feel better.	x	x	
30. When I'm upset, I start to feel very bad about myself.	x		
31. When I'm upset, I believe that wallowing in it is all that I can do.			x
35. When I'm upset, it takes me a long time to feel better.		x	
36. When I'm upset, my emotions feel overwhelming.	x		
Clarity Subscale			
*1. I am clear about my feelings.			
4. I have no idea how I am feeling.		x	x
5. I have difficulty making sense out of my feelings.	x	x	x
*7. I know exactly how I am feeling.			
9. I am confused about how I feel.	x	x	x

*Reverse scored items

participants completed an online battery of self-report measures. A total of 1360 participants began the survey. Of these, only those who completed all survey items and failed none of

the two embedded attention checks were retained, resulting in a total sample of 1181. Participants were largely female ($n = 836$, 70.79%) with a mean age of 20.7 ($SD = 5.15$). The ethnic

breakdown of the sample is as follows: Caucasian ($n = 675$, 57.15%), Multi-ethnic ($n = 191$, 16.17%), Hispanic/Latino ($n = 143$, 12.11%), Asian ($n = 90$, 7.62%), African American ($n = 49$, 4.15%), Native American ($n = 13$, 1.10%), Pacific Islander ($n = 6$, 0.51%), and Other ($n = 14$, 1.19%). The university institutional review board approved all procedures.

Measures

Demographics Questionnaire A questionnaire was administered assessing common demographic variables such as age, gender, and ethnicity.

Emotion Regulation Participants completed the DERS-36. Each of the three short forms of the DERS were computed from responses to the original DERS; participants did not complete short forms separately or respond to any DERS items twice, in order to avoid undue participant burden or fatigue. Items contained within each brief DERS are displayed in Table 1. The first short form, the DERS-16 (Bjureberg et al. 2016), was developed using highest item-total correlations from the original DERS-36 study, with items removed based on content validity judgment by the DERS-36 scale developer. The DERS-16 retained no items from the awareness subscale. The second short form, the DERS-SF (Kaufman et al. 2016), retained items based on confirmatory factor analysis of multiple published exploratory factor analysis (EFA) results, using three metrics developed to select items with the strongest factor loading on the primary scale and minimal cross-loadings across EFAs. The DERS-SF consists of 18 items, three of which load onto each of the original six subscales. The third short form, the DERS-18 (Victor and Klonsky 2016), was developed by selecting the three highest loading items on each of the original six subscales using factor loadings from the original DERS-36 study, thus retaining 18 items. A more detailed summary of the initial development of each short form is available in Online Resource 1. Cronbach's α for all versions of the DERS and their corresponding subscales are presented in Table 2.

Psychopathology

Several aspects of psychopathology shown to be associated with emotion regulation difficulties were assessed to examine the concurrent validity of the three brief DERS. Measures were selected in order to capture the same constructs as validity measures used in initial development of all three brief DERS (Bjureberg et al. 2016; Kaufman et al. 2016; Victor and Klonsky 2016).

Borderline Personality Disorder Symptoms Features of BPD were measured in the current sample using the short version of the Borderline Symptom List (BSL-23; Bohus et al. 2009). The BSL-23 measures a range of clinical symptoms associated with BPD over the course of the past week on a 0 (“not at all”) to 4 (“very strong”) Likert scale. This measure demonstrates good reliability and ability to discriminate BPD from less severe forms of psychopathology (Bohus et al. 2009), has been used to measure BPD in non-clinical samples with good reliability and validity (Salsman and Linehan 2012), and shows strong convergent validity with other measures of BPD (Glenn et al. 2009). In the current sample, Cronbach's α for this measure was 0.95.

Depression and Anxiety For a subset of the sample ($n = 576$), the 21-item version of the Depression Anxiety Stress Scales (DASS-21; Henry and Crawford 2005) was administered. This measure captures depression, anxiety, and stress experienced in the past week on a Likert scale ranging from 0 (“never”) to 4 (“most of the time”). In the current study, only the more pathological subscales of depression and anxiety were examined to reduce the number of comparisons. The DASS-21 has demonstrated good reliability, construct, and discriminant validity in a non-clinical sample (Henry and Crawford 2005) and was used in the development of the DERS-16 (Bjureberg et al. 2016). In the current study, Cronbach's α was 0.91 for depression and 0.80 for anxiety.

Clinically-Relevant Behaviors For the full sample, several problem behaviors associated with deficits in emotion

Table 2 Internal consistency of all DERS scales

	Cronbach's alpha			
	DERS-36	DERS-16	DERS-SF	DERS-18
Total scale	0.95	0.93	0.90	0.90
Nonacceptance	0.93	–	0.89	0.91
Goals	0.89	–	0.88	0.88
Impulse	0.87	–	0.89	0.89
Awareness	0.84	–	0.78	0.77
Strategies	0.91	–	0.84	0.84
Clarity	0.82	–	0.82	0.82

regulation were assessed using the behavioral supplement of the BSL-23. The 11-item supplement measures the frequency of occurrence of a variety of impulsive and harmful behaviors over the past week on a 0 (“not at all”) to 4 (“daily or more often”) Likert scale. Dichotomous variables were created for items capturing suicidality (e.g., “I told other people that I was going to kill myself”), self-harm (e.g., “I hurt myself by cutting, burning, strangling, headbanging, etc.”), disordered eating (e.g., “I had episodes of binge eating”), and substance use (e.g., “I got drunk”) such that any frequency of the behavior (rating of 1–4) was coded as positive for the behavior and zero ratings of the item corresponding to the behavior was coded as negative. The BSL supplement was used in the development of the DERS-16 (Bjureberg et al. 2016).

Data Analytic Plan

First, Cronbach’s alpha evaluated the internal consistency of all three short form total scores, as well as the original DERS-36. Next, four sets of regression analyses (12 total) examined the concurrent validity of each short form and the original DERS-36 total score on continuous variables of BPD, depression and anxiety symptoms, and four sets of logistic regression analyses (16 total) examined the concurrent validity of each short form and the original DERS-36 total score on dichotomous variables of suicidality, self-harm, disordered eating, and substance use. We did not include any covariates in the models. Analyses were conducted separately for each brief version of the DERS, due to highly overlapping items comprising each scale, leading to problematic multicollinearity if each were included within the same model ($r_s > .95$, $p_s < .001$). Bonferroni corrections at $p < .002$ were used to adjust for multiple comparisons for these analyses. We conducted post-hoc tests for differences between measures in the strength of relationships between DERS total scales and outcome measures with Dunn and Clark’s (1969) z test using the *cocor* package in R (Diedenhofen and Musch 2015).

Next, the internal consistency and concurrent validity of each DERS subscale was examined for the two short form measures that retained subscales (DERS-SF and DERS-18). To evaluate concurrent validity, we conducted three sets of multiple regression analyses (nine total multiple regressions) for the two short forms and original DERS-36 on continuous variables of BPD, depression, and anxiety, and three sets of multiple logistic regression analyses (12 total multiple logistic regressions) for the two short forms and original DERS-36 on dichotomous variables of suicidality, self-harm, disordered eating, and substance use. Given the large number of comparisons between DERS original and short form subscales, we plotted these results using the *arm* package in R (Gelman & Su 2016) for ease of interpretation. Bonferroni corrections at $p < .0004$ were used to adjust for multiple comparisons for these analyses.

Results

Reliability

Table 2 summarizes Cronbach’s alpha analyses. The total scores of all three short forms of the DERS demonstrated strong internal consistency reliability, comparable to each other and the original DERS-36. Similarly, internal consistencies for the DERS-SF and DERS-18 subscales were high and comparable to the original DERS-36 subscales. Additionally, consistent with previous research on the DERS-36, the awareness subscale of both short forms had the lowest reliability.

Concurrent Validity

Total Scores Regression analyses evaluating concurrent validity of the DERS-16, DERS-SF, and DERS-18 total scores on psychopathology and clinically-relevant behaviors are summarized in Table 3. Similar to the original DERS-36, all three short form versions of the DERS demonstrated strong associations with BPD symptoms, depression, anxiety, suicidality, self-harm, and disordered eating ($p_s < .001$). Substance use was not significantly associated with any of the DERS forms after adjusting for multiple comparisons using Bonferroni corrections ($p_s = .005$ to $.01$). Results from the z test indicated that total scores across all three brief versions demonstrated similar concurrent validity ($p_s > .05$). However, there was some evidence that short forms were more strongly associated with BPD and depression symptoms than the original DERS-36.

Subscale Scores Regression analyses evaluating concurrent validity of the DERS-SF and DERS-18 subscales, compared to the DERS-36 subscales, on psychopathology are summarized in Fig. 1. All three measures demonstrated a high degree of concordance in their pattern of association with clinical variables. Comparable to the original DERS-36 subscales, subscales on both the DERS-SF and DERS-18 demonstrated strong associations with BPD symptoms, depression, and anxiety with similar patterns of significance in individual subscales. The strategies and clarity subscales were associated with BPD symptoms across the DERS-36, DERS-SF, and DERS-18 ($p_s < .0001$). However, the nonacceptance subscale was also significantly associated with BPD symptoms on the DERS-SF and DERS-18 ($p_s < .0001$). The strategies and clarity subscales of the DERS-36 were associated with depression ($p_s < .0001$). Similarly, the strategies and clarity subscales were associated with depression for both short forms ($p_s < .0001$); however awareness, for both short forms, was also associated with depressive symptoms ($p_s < .0003$).

Table 3 Multiple linear and binary logistic regressions assessing the predictive validity of the DERS-36 and multiple short form total scores on all outcomes

Predictor Variables and Outcomes	DERS-36	DERS-16	DERS-SF	DERS-18
Multiple linear regression: b(s.e.)				
BPD	.45(.01)* _a	.79(.03)* _{a,b}	.85(.03)* _b	.84(.03)* _b
Depression	.13(.01)* _a	.21(.01)* _b	.24(.01)* _{a,b}	.24(.01)* _{a,b}
Anxiety	.09(.01)* _a	.16(.01)* _a	.17(.01)* _a	.17(.01)* _a
Binary logistic regression: OR(s.e.)				
Suicidality	1.03(.01)* _a	1.04(.01)* _a	1.05(.01)* _a	1.05(.01)* _a
Self-harm	1.04(.01)* _a	1.06(.01)* _a	1.07(.01)* _a	1.07(.01)* _a
Disordered Eating	1.02(.003)* _a	1.04(.01)* _a	1.05(.01)* _a	1.05(.01)* _a
Substance Use	1.01(.002) _a	1.01(.004) _a	1.01(.01) _a	1.01(.01)* _a

*Significant after Bonferonni corrections at $p < .002$. Different subscripts per row indicate significant differences between DERS forms using Dunn and Clark' (1969) z test

The pattern of subscales associated with anxiety was identical across all measures: the strategies and clarity subscales of the DERS-36, DERS-18, and DERS-SF were associated with anxiety ($ps < .0002$).

Binary logistic regression analyses evaluating concurrent validity of the DERS-SF and DERS-18 subscales on clinically relevant behaviors are summarized in Fig. 2. Few significant associations were detected between any of the DERS subscale

scores and binary outcomes; however, the DERS-SF and DERS-18 subscales demonstrated nearly identical patterns of concurrent validity. Although all versions of the DERS demonstrated a similar pattern of effects, only the goals subscale of the DERS-18 was significantly associated with suicidality ($p = .0002$), and only the clarity subscale of the DERS-36 was associated with disordered eating ($p = .0003$) after correcting for multiple comparisons.

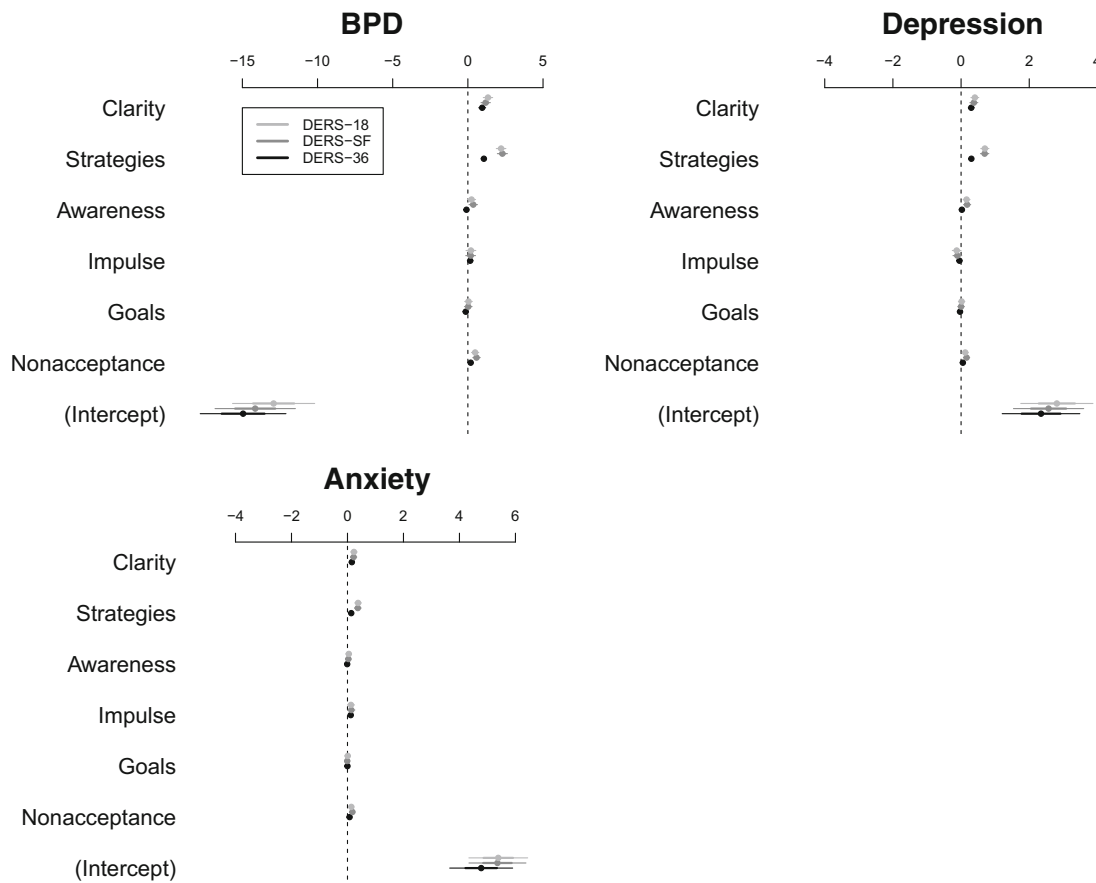


Fig. 1 Plots of multiple linear regression unstandardized coefficients

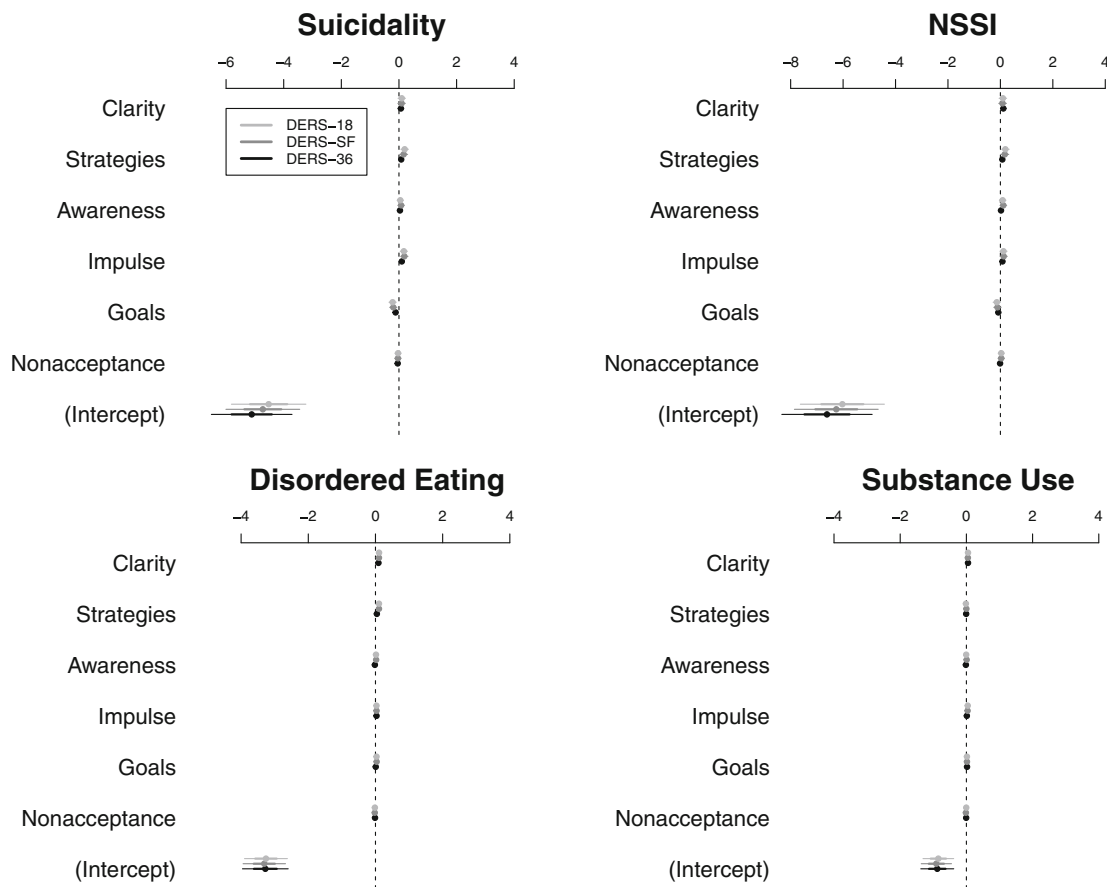


Fig. 2 Plots of multiple logistic regression unstandardized coefficients

Discussion

The current study evaluated three new brief versions of the widely-used DERS for both internal consistency reliability and concurrent validity with relevant clinical outcomes to guide measure selection for future research. Total scores for all three short forms demonstrated similar reliability and concurrent validity for BPD, depression and anxiety symptoms, and a range of clinically relevant problem behaviors. There was slight evidence that the short forms were more strongly associated with BPD and depression than the DERS-36, lending further support for the utility of employing a brief version. Taken together, the results suggest that all three brief versions performed equally well and are highly suited for the evaluation of emotion regulation difficulties in research settings, especially those in which brevity is ideal due to time constraints.

The results of this study do not clearly answer the question of which brief form of the DERS is the best to use in clinical and research settings. All three versions were well-developed and tested to ensure sound psychometric properties (Bjureberg et al. 2016; Kaufman et al. 2016; Victor and Klonsky 2016; Online Resource 1) and the current study supported these

initial findings in an independent sample. Therefore, one interesting conclusion from this study is that different psychometrically sound and comparably performing brief instruments can be derived from the same measure using discrepant methods. Unfortunately, the development of three separate, but equally reliable and valid short versions of the DERS is likely to result in measurement inconsistency across studies, limiting the ability to compare or combine results across different research groups and potentially slowing the research process. This specific example reflects a larger concern related to the proliferation of similar measures within psychology generally, and within the field of emotion regulation specifically. The precise manner of defining and measuring emotion regulation, like many psychological constructs, has been the source of much debate (e.g., Cole et al. 2004; Gross and Barrett 2011; Seligowski and Orcutt 2015). Therefore, different research groups have designed varying measures to try to most effectively capture the construct; however, discrepant research findings across studies can be difficult to interpret when different assessment approaches are used (e.g., Naragon-Gainey et al. 2017). The ability to compare findings across studies is crucial to the advancement of an empirical science; therefore, lack of consistency and redundancy across

multiple measures of the same construct makes clarification of research findings more difficult (Carlson and Herdman 2012; Bloch et al. 2010). Alliance and collaboration among emotion regulation researchers in measure development might help the field move forward with clearer construct conceptualizations and more comparable research methods. Along these lines, we would strongly recommend that experts within the emotion regulation field select one short version of the DERS that is believed to be best suited to briefly assess emotion regulation and proceed with utilizing this measure consistently across studies. Although it is certainly difficult to derive expert consensus on the recommended measurement of a construct, recent examples exist within psychology in which this has been done effectively (Bolte et al. 2018; Howes et al. 2018).

Although our findings did not clearly identify a brief version of the DERS with the best psychometrics, they did provide some clues as to which of these measures might provide optimal utility. Subscale analyses lent support for the utility of selecting a short form that retains the ability to examine specific facets of emotion regulation. As the DERS-36 has demonstrated in several research studies (e.g., Buckholdt et al. 2015; Mennin et al. 2009; Weinbach et al. 2017), specific subscales of the DERS-SF and the DERS-18 were associated with different clinical problems. For example, although the strategies and clarity subscales were associated with a number of different clinical outcomes across measures, the awareness subscale was especially associated with depression, and the nonacceptance subscale with BPD symptoms for both the DERS-18 and DERS-SF. As highlighted in this example, as compared to the DERS-36, the DERS-SF and DERS-18 sometimes offered a more specific subscale profile for different clinical issues (e.g., BPD symptoms, depression, suicidality). While this pattern certainly warrants replication, and research testing the predictive validity of the short forms is needed, it does offer increased utility of the measures. In addition, gold standard practices for developing short form versions of assessment measures mandate that all facets captured by the original measure be preserved to the degree possible (Smith et al. 2000). This increased level of specificity and content overlap between short and long forms is useful for better understanding the core emotion regulation mechanisms underlying problem behaviors, may support the assessment of certain clinical problems based on specific emotion regulation difficulties, and would inform more effective intervention and efficient treatment targets for specific clinical problems.

Interestingly, although the utility of the awareness subscale has been questioned due to concerns about construct validity and previous analyses suggesting that it is less related to the overall DERS score (Bardeen et al. 2012), the DERS-16—which removed all items from the awareness subscale—did not result in better concurrent validity compared to brief measures which retained items from the awareness subscales. Moreover, the awareness subscale of

both the DERS-SF and the DERS-18 was significantly associated with depression, lending support for the retention of that subscale. Although the DERS-16 developers omitted the awareness subscale for theoretically-based reasons and in response to lower reliability scores across studies, this elimination may result in a less valid measure that does not maintain all important constructs captured with the DERS-36 (Smith et al. 2000). In sum, these findings suggest that the DERS-SF and DERS-18 may have some incremental advantage over the DERS-16. Although all the brief scales are newly developed, and none appear to be widely used as of yet, the DERS-SF has been in use longer and cited more often than the DERS-18, which suggests that there could be utility in continuing to use this measure going forward in order to compare results between the greatest number of studies. However, further work is needed investigating which DERS subscales might be most useful for research on emotion regulation. Furthermore, findings from the current study suggest that factor analysis of the DERS-16 to identify subscales could increase its utility.

The strengths of the current study include the use of a large sample, which successfully provides further validation of the three brief DERS scales by replicating initial reliability and concurrent validity analyses. Limitations include the use of cross-sectional self-report in a non-clinical, undergraduate sample which is largely female and Caucasian. Further testing in a more diverse sample is warranted. In addition, while measures were selected to capture the same constructs used in initial validation of all three brief DERS, the current study's measures overlapped with some (i.e., DASS-21, BSL) but not all measures used in the original studies (e.g., DSHI, BEST, BDI-II), which may have impacted the results (Bjureberg et al. 2016; Kaufman et al. 2016; Victor and Klonsky 2016). Moreover, in the current study and in all the initial brief DERS developmental studies but one (Study 1; Bjureberg et al. 2016), the short versions are derived from the DERS-36. Further research should examine potential differences in response styles if the short versions are administered alone. Given the nature of emotion regulation and the strong association between DERS scores and clinical variables, further testing should also be done in a clinical sample. The current study also did not evaluate the test-retest reliability of the short forms. Given that briefer versions are ideal when repeated measurement is needed, future studies should examine their utility in this domain. Similarly, in clinical settings, it is important to consider measurement precision for shortened measures (Kruyen et al. 2013). Future studies should examine how these three measures compare on measurement precision. Additionally, future longitudinal studies are warranted to establish the long-term psychometrics, including predictive validity, of these measures.

Conclusions

This study is the first to provide additional support for the psychometrics and potential utility of all three existing brief DERS. Although the results of the current study did not clearly identify one brief version of the DERS as superior to the others, they did offer some clarification regarding potential important differences between multiple short forms of this measure. In particular, our results provide support for the retention and use of subscales of the DERS to increase the utility of the instrument in association with specific clinical outcomes. We hope that unified efforts to clarify the definition and measurement of emotion regulation will continue in order to enhance the validity, consistency, and comparability of the assessment of this important construct.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Experiment Participants The university Institutional Review Board (IRB) approved and oversaw all study procedures, including informed consent.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

- Banyard, V., Hamby, S., & Grych, J. (2017). Health effects of adverse childhood events: Identifying promising protective factors at the intersection of mental and physical well-being. *Child Abuse & Neglect*, *65*, 88–98.
- Bardeen, J. R., Fergus, T. A., & Orcutt, H. K. (2012). An examination of the latent structure of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, *34*, 382–392.
- Bjureberg, J., Ljotsson, B., Tull, M. T., Hedman, E., Sahlin, H., Lundh, L., et al. (2016). Development and validation of a brief version of the difficulties in emotion regulation scale: The DERS-16. *Journal of Psychopathology and Behavioral Assessment*, *38*(2), 284–296.
- Bloch, L., Moran, E. K., & Kring, A. M. (2010). On the need for conceptual and definitional clarity in emotion regulation research on psychopathology. In A. M. Kring & D. M. Sloan (Eds.), *Emotion regulation and psychopathology: A transdiagnostic approach to etiology and treatment* (pp. 88–107). New York: Guilford Press.
- Bohus, M., Kleindienst, N., Limberger, M. F., Stieglitz, R. D., Domsalla, M., Chapman, A. L., & Wolf, M. (2009). The short version of the borderline symptom list (BSL-23): Development and initial data on psychometric properties. *Psychopathology*, *42*, 32–39.
- Bolte, S., Mahdi, S., Gau, S. S., Holtmann, M., Karande, S., Rohde, L. A., et al. (2018). Standardized assessment of functioning in ADHD: Consensus on the icf core sets for ADHD. *European Child & Adolescent Psychiatry*. <https://doi.org/10.1007/s00787-018-1119-y>.
- Brockmeyer, T., Skinde, M., Wu, M., Breelein, E., Rudofsky, G., Herzog, W., & Friederich, H. C. (2014). Difficulties in emotion regulation across the spectrum of eating disorders. *Comprehensive Psychiatry*, *55*(3), 565–571.
- Buckholdt, K. E., Parra, G. R., Anestis, M. D., Lavender, J. M., Jobe-Shields, L. E., Tull, M. T., & Gratz, K. L. (2015). Emotion regulation difficulties and maladaptive behaviors: Examination of deliberate self-harm, disordered eating, and substance misuse in two samples. *Cognitive Therapy and Research*, *39*(2), 140–152.
- Carlson, K. D., & Herdman, A. O. (2012). Understanding the impact of convergent validity on research results. *Organizational Research Methods*, *15*(1), 17–32.
- Cicchetti, D., Ackerman, B. P., & Izard, C. E. (1995). Emotions and emotion regulation in developmental psychopathology. *Development and Psychopathology*, *7*, 1–10.
- Cole, P. M., Martin, S. E., & Dennis, T. A. (2004). Emotion regulation as a scientific construct: Methodological challenges and directions for child development research. *Child Development*, *75*(2), 317–333.
- Crowell, S. E., Beauchaine, T. P., Hsiao, R. C., Vasilev, C. A., Yaptangco, M., Linehan, M. M., & McCauley, E. (2012). Differentiating adolescent self-injury from adolescent depression: Possible implications for borderline personality development. *Journal of Abnormal Child Psychology*, *40*, 45–57.
- Diedenhofen, B., & Musch, J. (2015). Cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS One*, *10*(4), e0121945.
- Dunn, O. J., & Clark, V. (1969). Correlation coefficients measured on the same individuals. *Journal of the American Statistical Association*, *64*(325), 366–377.
- Ford, B. Q., Mauss, I. B., Troy, A. S., & Hankin, B. (2014). Emotion regulation moderates the risk associated with the 5-HTT gene and stress in children. *Emotion*, *14*(5), 930–939.
- 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. *Journal of Psychiatric Research*, *58*, 175–180.
- Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, *73*(2), 349–360.
- Garnefski, N., Kraaij, V., & Spinhoven, P. (2001). Negative life events, cognitive emotion regulation and emotional problems. *Personality and Individual Differences*, *30*, 1311–1327.
- Gelman, A., Su, Y. S. (2018). Arm: Data analysis using regression and multilevel/hierarchical models. R package version 1.10-1. <https://CRAN.R-project.org/package=arm>. Accessed 10 May 2018
- Glenn, C. R., Weinberg, A., & Klonsky, E. D. (2009). Relationship of the borderline symptom list to DSM-IV borderline personality disorder criteria assessed by semi-structured interview. *Psychopathology*, *42*, 394–398.
- 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. *Journal of Psychology and Behavioral Assessment*, *26*(1), 41–54.
- Gratz, K. L., & Tull, M. T. (2010). Emotion regulation as a mechanism of change in acceptance- and mindfulness-based treatments. In R. A. Baer (Ed.), *Assessing mindfulness and acceptance: Illuminating the processes of change* (pp. 107–134). Oakland: New Harbinger Publications.
- Gross, J. J., & Barrett, L. F. (2011). Emotion generation and emotion regulation: One or two depends on your point of view. *Emotion Review*, *3*, 8–16.
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect,

- relationships, and well-being. *Journal of Personality and Social Psychology*, 85(2), 348–362.
- Henry, J. D., & Crawford, J. R. (2005). The short-form version of the depression anxiety stress scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 44, 227–239.
- Hofmann, S. G., Carpenter, J. K., & Curtiss, J. (2016). Interpersonal emotion regulation questionnaire (IERQ): Scale development and psychometric characteristics. *Cognitive Therapy and Research*, 40(3), 341–356.
- Howes, O. D., Rogdaki, M., Findon, J. L., Wichers, R. H., King, B. H., Loth, E., et al. (2018). Autism spectrum disorder: Consensus guidelines on assessment, treatment and research from the British Association for Psychopharmacology. *Journal of Psychopharmacology*, 32, 3–29.
- Hutchison, J. A., & Gunther, K. C. (2013). Measuring beliefs about automatic mood regulation: Development of a self-report scale. *Cognitive Therapy and Research*, 37, 1243–1256.
- Joomann, J., & Siemer, M. (2014). Emotion regulation in mood disorders. In J. J. Gross (Ed.), *Handbook of emotion regulation, second edition* (pp. 413–427). New York: Guilford Press.
- 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. *Journal of Psychopathology and Behavioral Assessment*, 38(3), 443–455.
- Kruyen, P. M., Emons, W. H. M., & Sijtsma, K. (2013). Shortening the S-STAI: Consequences for research and clinical practice. *Journal of Psychosomatic Research*, 75, 167–172.
- Linehan, M. M. (1993). *Cognitive behavioral therapy of borderline personality disorder*. New York: Guilford Press.
- Mennin, D. S., McLaughlin, K. A., & Flanagan, T. J. (2009). Emotion regulation deficits in generalized anxiety disorder, social anxiety disorder, and their co-occurrence. *Journal of Anxiety Disorders*, 23(7), 866–871.
- Naragon-Gainey, K., McMahon, T. P., & Chacko, T. P. (2017). Structure of common emotion regulation strategies: A meta-analytic examination. *Psychological Bulletin*, 143(4), 384–427.
- Neacsiu, A. D., Fang, C. M., Rodriguez, M., & Rosenthal, Z. M. (2017). Suicidal behavior and problems with emotion regulation. *Suicide and Life-threatening Behavior*. <https://doi.org/10.1111/sltb.12335>.
- Neumann, A., van Lier, P. A. C., Gratz, K. L., & Koot, H. M. (2010). Multidimensional assessment of emotion regulation difficulties in adolescents using the difficulties in emotion regulation scale. *Assessment*, 17, 138–149.
- Niven, K., Totterdell, P., Stride, C. B., & Holman, D. (2011). Emotion regulation of others and the self (EROS): The development and validation of a new individual difference measure. *Current Psychology*, 30, 50–73.
- Osborne, T. L., Michonski, J., Saysrs, J., Welch, S. S., & Anderson, L. K. (2017). Factor structure of the difficulties in emotion regulation scale (DERS) in adult outpatients receiving dialectical behavior therapy (DBT). *Journal of Psychopathology and Behavioral Assessment*, 39, 355–371.
- Salsman, N. L., & Linehan, M. M. (2012). An investigation of the relationships among negative affect, difficulties in emotion regulation, and features of borderline personality disorder. *Journal of Psychopathology and Behavioral Assessment*, 34(2), 260–267.
- Seligowski, A. V., & Orcutt, H. K. (2015). Examining the structure of emotion regulation: A factor-analytic approach. *Journal of Clinical Psychology*, 0, 1–19.
- Sharp, C., Venta, A., Vanwoerden, S., Schramm, A., Ha, C., Newlin, E., et al. (2016). First empirical evaluation of the link between attachment, social cognition and borderline features in adolescents. *Comprehensive Psychiatry*, 64, 4–11.
- Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12, 102–111.
- Thompson, R. A. (1994). Emotion regulation: A theme in search of definition. *Monographs of the Society for Research in Child Development*, 59(2–3), 25–52.
- Victor, S. E., & Klonsky, E. D. (2016). Validation of a brief version of the difficulties in emotion regulation scale in five samples. *Journal of Psychopathology and Behavioral Assessment*, 38(4), 582–589.
- Weinbach, N., Sher, H., & Bohon, C. J. (2017). Differences in emotion regulation difficulties across types of eating disorders during adolescence. *Journal of Abnormal Child Psychology*. <https://doi.org/10.1007/s10802-017-0365-7>.
- Weiss, N. H., Bold, K. W., Sullivan, T. P., & Armeli, S. (2017). Testing bidirectional associations among emotion regulation strategies and substance use: A daily diary study. *Addiction*, 112(4), 695–704.
- Wilks, C. R., Korslund, K. E., Harned, M. S., & Linehan, M. M. (2016). Dialectical behavior therapy and domains of functioning over two years. *Behaviour Research and Therapy*, 77, 162–169.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.