

Validation of a Brief Version of the Difficulties in Emotion Regulation Scale (DERS-18) in Five Samples

Sarah E. Victor¹ · E. David Klonsky¹

© Springer Science+Business Media New York 2016

Abstract The field of affective clinical science has expanded dramatically over the past several decades (Rottenberg and Gross 2003; Tracy et al. 2014). An important part of this research is understanding emotion regulation and dysregulation, in particular, how individuals differ in their ability to identify, accept, and manage their emotional experiences. One of the most common and widely cited measures of emotion dysregulation is the Difficulties with Emotion Regulation Scale (DERS; Gratz and Roemer 2004), which has been extensively used to facilitate understanding of how emotion dysregulation is associated with psychiatric symptoms, other emotion-related constructs, and treatment progress. While this scale has contributed greatly to our understanding of emotion regulation problems, its length makes its inclusion in brief study protocols difficult, limiting its utility and increasing participant burden. In order to address this issue, we developed a short form version of the DERS (DERS-18) composed of the strongest items from each of the measure's six subscales from the original DERS publication (Gratz and Roemer 2004), and then validated this measure in five datasets that vary in age and sample type. Our results demonstrate that an 18-item short-form of the DERS exhibits a similar structure as the original 36-item DERS, demonstrates excellent reliability and validity, and performs similarly to the original DERS despite comprising half the items. The DERS-18 has the potential to improve and expand emotion regulation assessment while reducing demands on research participants.

Keywords Emotion · Emotion regulation · Measurement

Extensive research has demonstrated the importance of understanding how emotions influence psychopathology (Gross and Jazaieri 2014). In particular, understanding how people regulate their emotions has become a major area of focus for affective scientists (Tracy et al. 2014), with researchers using emotion regulation and dysregulation to explain important clinical phenomena such as depression (Joormann and Vanderlind 2014), anxiety (Mennin et al. 2009), and Borderline Personality Disorder (Glenn and Klonsky 2009).

Emotion regulation has been broadly defined as “the processes by which individuals influence which emotions they have, when they have them, and how they experience and express these emotions” (Rottenberg and Gross 2003, p. 229). Gratz and Roemer (2004) expanded on this idea by considering the components necessary for successful emotion regulation, including awareness and acceptance of one's emotions, the ability to change emotions in accordance with one's goals, and the ability to control behavior in the face of negative emotions, with emotion dysregulation being the absence of any of these abilities (p. 42–43). This multidimensional definition has been used by researchers in clinical psychology (Donahue et al. 2014; Racine and Wildes 2013) to understand how individuals differ not only in overall emotion regulation capacity, but also in specific emotion regulation deficits.

The Difficulties in Emotion Regulation Scale (DERS; Gratz and Roemer 2004) was developed to facilitate thorough assessment of emotion dysregulation in a variety of populations. The DERS is a 36-item self-report questionnaire measuring emotion dysregulation that includes six factor-analytically-derived subscales: lack of awareness of one's emotions (awareness), lack of clarity about the nature of one's emotions (clarity), lack of acceptance of one's emotions (nonacceptance), lack of access to effective emotion regulation strategies (strategies), lack

✉ Sarah E. Victor
sevictor@psych.ubc.ca

¹ Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC V6T1Z4, Canada

of ability to engage in goal-directed activities during negative emotions (goals), and lack of ability to manage one's impulses during negative emotions (impulse) (Gratz and Roemer 2004). These subscales and the DERS total score show very high internal consistency and show strong convergent validity with other measures of related constructs, such as experiential avoidance, emotional expressivity, and internalizing and externalizing problems.

The DERS has been cited hundreds of times and has been used in a wide range of study designs, populations of interest, and types of psychopathology being studied. The DERS was originally validated on a sample of college undergraduates, but it has since been used in studies of adolescents (Weinberg and Klonsky 2009), adults (Bardeen and Fergus 2014), and older adults (Staples and Mohlman 2012). It has also been used in studies ranging from large samples of nonclinical populations (Sharp et al. 2014) to patients with severe psychopathology (Fowler et al. 2014). It is most often used cross-sectionally as a measure of trait-like emotion dysregulation (e.g., Franklin et al. 2010), but it has also been used to predict psychopathology longitudinally (e.g., Scott et al. 2014) and as a measure of clinical outcome in treatment trials (e.g., Gratz et al. 2014).

Extensive research supports the reliability and validity of the DERS, including work establishing the psychometric properties of the measure in different languages and populations. Research supports the DERS's test-retest reliability (Gratz and Roemer 2004; Staples and Mohlman 2012), internal consistency (Weinberg and Klonsky 2009), and convergent and divergent validity with measures of other psychological constructs (Bardeen and Fergus 2014). There remains, however, some debate regarding the most appropriate factor structure for the measure. While some authors have found evidence to support the original, six-factor structure of the DERS (Giromini et al. 2012; Mitsopoulou et al. 2013; Ruganci and Gençöz 2010), others have suggested that a five-factor model may better fit the data (Bardeen et al. 2012; Cho and Hong 2013). In one study, confirmatory factor analysis in a sample of undergraduate women suggested the DERS Awareness subscale did not represent the same superordinate emotional regulation factor as the remaining 5 subscales, leading the authors to argue for this subscale's removal from the DERS (Bardeen et al. 2012); these authors did acknowledge, however, that the Awareness subscale may have utility in and of itself. In a sample of undergraduate students assessed with the Korean translation of the DERS, results supported combining the Clarity and Awareness subscales into a single subscale (Understanding; Cho and Hong 2013). In at least one study, there was equivalent support for a five- and six-factor structure, with the authors arguing for a six-factor structure for consistency with published literature (Fowler et al. 2014).

While research has generally supported the psychometric properties of the DERS, some work suggests that the measure

is unnecessarily long. In a large sample of Greek adults, for example, a 30-item, six-factor scale was a better fit to the data than the 36-item version (Mitsopoulou et al. 2013). Thirty-item scales have been validated in other populations as well (Cooper et al. 2014; Kökönyei et al. 2014), although in all cases, the length of the measure was not the primary focus of the analyses.

A common issue in research protocol development is maintaining a balance between the use of valid and reliable measures, regardless of length, while avoiding unnecessary burden on participants. As a result, shortened versions of questionnaires are often developed with the dual goals of reducing participant burden and retaining the reliability and validity of the original measure. In addition to the financial and time-related costs of using excessively long measures, research also suggests that the use of longer questionnaire batteries can impact data collection and quality; for example, studies have shown that participants are less willing to begin and complete longer questionnaire batteries compared to shorter ones (Roszkowski and Bean 1990), and that response quality declines as questionnaire duration increases (Galesic and Bosnjak 2009).

To address these concerns, we developed a short-form of the DERS (DERS-18¹) that preserved its reliability, validity, and factor structure while reducing participant research burden. While we appreciate the existing debate regarding the most appropriate factor structure for the DERS, and the Awareness subscale in particular (see, for example, Bardeen et al. 2012; Weinberg and Klonsky 2009), our aim was relatively specific: to develop a shorter version that best mirrors the existing, full-length DERS. We therefore chose to retain items representing each of the original six factors. To do this, we created a half-length version of the scale using original factor loadings provided by Gratz and Roemer (2004), and then validated this new, shorter version in five varied samples.

Methods

Participants and Data Collection Procedures

Participants completed the original 36-item DERS within a battery of self-report questionnaires asking about a variety of psychological constructs, such as emotional experiences, personality, non-suicidal self-injury, and behavioral difficulties. The subset of items used for the DERS-18 were taken from the full, original version of the DERS; participants did not complete the DERS-18 separately and they did not respond to any DERS items twice. All studies were conducted in accordance

¹ While this manuscript was under review, another short form of the DERS was published (DERS-SF; Kaufman et al. 2015). We, therefore, refer to our measure as the DERS-18 in order to distinguish the two novel forms of the DERS, which differ slightly in item selection, validation procedures, and sample types.

with the ethics review board at the university housing the principal investigator. All adults provided informed consent, and all adolescents provided informed assent, prior to participation; for adolescents, informed consent from a parent or guardian was also obtained.

Sample 1 ($n = 429$) was drawn from high school students located in the northeastern United States (for previously reported data using this sample, see Klonsky et al. 2013). Participants ranged in age from 13 to 17 and were primarily female ($n = 265$, 61.45 %). The study was completed using paper-and-pencil questionnaires. While we have published psychometric data on the DERS in this sample previously (Weinberg and Klonsky 2009), those findings were using the original, long form version, and we have not previously analyzed or reported findings utilizing the proposed short form.

Sample 2 ($n = 167$) was drawn from adolescents receiving inpatient psychiatric care at a facility in the northeastern United States (for previously reported data using this sample, see Glenn and Klonsky 2013; Klonsky et al. 2013). Participants were primarily female ($n = 129$, 77.25 %) and Caucasian ($n = 112$, 67.07 %), with a mean age of 15.61 ($SD = 1.42$). Participants were recruited for participation and completed the questionnaires in person, using paper-and-pencil measures. In addition to university ethics approval, this study was also approved by and conducted in accordance with the Institutional Review Board at South Oaks Hospital, where data were collected.

Sample 3 ($n = 160$) was comprised of adults drawn from a university community in western Canada for a study on emotional experiences and non-suicidal self-injury. The sample was primarily comprised of undergraduate students enrolled in the study for course credit, with individuals reporting a history of self-injury over-represented in the study sample (for previously reported data using this sample, see Victor and Klonsky 2014). Participants completed computerized questionnaires alone in a laboratory setting. Participants were primarily female ($n = 109$, 68.13 %) and Asian descent ($n = 106$, 67.09 %), with a mean age of 23.28 ($SD = 5.45$).

Sample 4 ($n = 163$) was recruited as part of the same study on emotional experiences in non-suicidal self-injury as sample 3, but was comprised of adults recruited using an online system (Amazon.com's Mechanical Turk). Within the Mechanical Turk system, potential study participants can view information about a study's eligibility, procedures, and compensation, and can answer surveys and questionnaires from home. Research suggests that Mechanical Turk can be a valuable source for easy to obtain, reliable self-report data (Buhrmester et al. 2011). Because the same eligibility criteria were used for both samples, adults with a recent history of self-injury are over-represented in this sample as well (for previously reported data using this sample, see Victor and Klonsky 2014). Participants were approximately half female ($n = 91$, 55.83 %), majority Caucasian ($n = 109$, 70.32 %), and on average 30.49 years old ($SD = 10.73$).

Sample 5 ($n = 705$) was recruited as part of a study evaluating the psychometric properties of a new measure of emotional experiences among adults from the community. Similar to sample 4, these participants were recruited using Amazon's Mechanical Turk. Because this study was not investigating non-suicidal self-injury, the sample more closely approximates a non-clinical, community sample than samples 3 or 4. Participants were primarily female ($n = 411$, 58.30 %) and Caucasian ($n = 516$, 75.44 %), with a mean age of 35.26 ($SD = 13.19$).

Scale Development and Scoring

To develop the DERS-18, we considered two methods for selection of appropriate items; first, we utilized the published factor loadings for the original 36 items from Gratz and Roemer (2004). These factor loadings were derived from an exploratory factor analysis, after which items with poor or dual loadings ($<.40$ on any subscale or $>.40$ on two or more subscales) were excluded. The subsequent factor analysis, which included the 36 items that comprise the DERS, resulted in six factors comprising between five and eight items per factor. In order to create a measure of half the length of the original, we selected the three highest-loading items on each factor (range: .61 to 1.00) for retention in the short form. Scoring for each item and factor was identical to that previously reported by Gratz and Roemer (2004), whereby each item is rated on a scale from 1 (almost never) to 5 (almost always), and scores for each subscale and total score are sums of the relevant items (with three of 18 items reverse-coded).

In order to clarify whether these items would perform similarly in our own data, we conducted an exploratory factor analysis using the same specifications as Gratz and Roemer (2004) in our own combined sample ($N = 1602$). Specifically, we used principal axis factoring with a fixed six-factor solution using promax oblique rotation with Kaiser normalization. The three highest-loading items on each factor in this analysis were identical to those supplied by Gratz and Roemer (2004) in 35 of 36 cases, with the remaining single item drawn from Gratz and Roemer's factor loadings (item #31, Strategies subscale) still loading strongly on its appropriate subscale in our own analysis (.62). None of the items loaded onto more than one factor (all other factor loadings $<.24$). Thus, given our aim to capture the existing DERS as closely as possible using a shorter version, we chose to retain the highest-loading items published by Gratz and Roemer (2004; full results of our factor analysis available from the authors on request).

Data Analyses

To evaluate whether the items selected for the shortened version retained the factor structure of the original version, using the same methodology as that reported by Gratz and Roemer (2004), we conducted an exploratory factor analysis with the

18 items using principal axis factoring and promax oblique rotation with Kaiser normalization. First, we conducted an analysis with a fixed six-factor solution in the combined samples (total $N = 1602$), and then used the same method in each of the five separate samples to confirm the factor structure in each case.

To determine whether the DERS-18 is internally consistent, Cronbach's alphas were calculated for each of the six subscale scores and the overall score for the DERS-18 in the combined sample. Subsequently, Cronbach's alphas were calculated for these same scores in each of the five samples separately, to confirm that the measure retains its reliability in samples varying by age and clinical severity. To examine the concurrent validity of the DERS-18, we calculated Pearson's correlation coefficients between each of the DERS-18 subscale scores and the overall score with their original, longer form counterpart. These correlations were evaluated for the combined sample as well as for each sample separately.

In addition, to assess convergent validity, we examined the association between the DERS-18 and measures of Borderline Personality Disorder (BPD) symptoms in each sample. Emotion dysregulation is a core feature of BPD (Glenn and Klonsky 2009), and the original DERS has exhibited robust correlations with BPD (Chapman et al. 2008; Scott et al. 2014). For samples 1, 3, 4, and 5, BPD was measured with the McLean Screening Inventory for BPD symptoms (MSI-BPD; Zanarini et al. 2003), a self-report measure of BPD, and for samples 2 and 3, BPD was measured using the BPD items from the Structured Interview for DSM-IV Personality (SIDP; Pfohl et al. 1997), a semi-structured interview assessing BPD symptoms.

We were also able to investigate predictive validity in sample 3. In that study, participants completed the DERS during a laboratory session and then subsequently completed a daily diary assessing positive and negative emotional experiences (for details regarding this portion of the assessment, see Victor and Klonsky 2014). For twenty specific emotions (10 positive, 10 negative), participants were asked once daily over a 2 week period the extent to which they experienced that emotion and, if they experienced the emotion at all, how difficult it was for them to regulate that emotion on a scale from 1 (very easy to regulate) to 5 (very difficult to regulate). We then created a "positive regulation" score by averaging the difficulty in regulating the 10 positive emotions across all diary entries, as well as a "negative regulation" score by averaging the difficulty in regulating the 10 negative emotions across all diary entries.

Results

Descriptive Characteristics and Factor Structure of the DERS-18

The DERS-18 is comprised of three items per subscale, for six subscales, or a total of 18 items. These include three reverse-

coded items, all on the Awareness subscale. In the combined sample, item total correlations ranged from .53 to .83, which was similar to those reported by Gratz and Roemer (2004) for the original DERS (.45 to .81). The mean inter-item correlations for each subscale and the overall score ranged from .53 to .75, which was slightly higher but similar to those reported for the original (.40 to .62).

As with the original DERS, all DERS-18 subscales were statistically significantly correlated with each other, with the exception of the Awareness and Goals subscales, which were weakly and non-significantly correlated. Correlations were generally strongest between the Strategies subscale and other subscales, with the exception of Awareness. Additional information on these correlations can be found in Table 1.

Our factor analysis, which used a fixed six-factor solution, accounted for 78.43 % of the variance. The DERS-18 items loaded onto their original, expected subscales (all loadings $> .75$), and no items loaded substantially onto any other subscale (all loadings $< .18$). This pattern was subsequently found in each of the separate samples. Factor loadings for the combined sample can be found in Table 2; factor loadings for each sample separately are available by request to the corresponding author.

Reliability and Validity of the DERS-18

One of the strengths of the original DERS is its high internal consistency reliability for the subscales and total score (Gratz and Roemer 2004). The DERS-18 retained very high internal consistency in the combined sample, with subscale alphas ranging from .77 (Awareness) to .90 (Goals and Impulse), and an overall score alpha of .91. In the five samples analyzed separately, generally excellent internal consistency was maintained, with most alphas greater than .80. In general, the Awareness subscale exhibited lower internal consistency, but remained in the acceptable to good range (Cronbach's alphas .69 to .80 across samples). Further details on the reliability of the DERS-18 in the combined and separated samples can be found in Table 3.

To examine concurrent validity, we computed correlations between the original DERS and the DERS-18. As expected, the DERS-18 subscale and total scores were strongly correlated with subscale and total scores on the original DERS. In the combined sample, correlations ranged from .92 (Awareness subscale) to .98 (total score). Across the separate samples, no short-form scale was correlated less than .90 with its longer equivalent, and the average correlation was .94. Additional information about DERS-18 validity can be found in Table 3.

To evaluate convergent validity of the DERS-18, we calculated correlations between the DERS-18 total score and measures of BPD in each of the five samples. For samples 1, 3, 4, and 5, correlations between the DERS-18 total score and MSI-BPD total score ranged from .49 (sample 4, community

Table 1 Correlations between subscales of the DERS-18 in the combined sample ($N = 1602$)

	Goals	Nonacceptance	Impulse	Clarity	Awareness	Strategies
Goals	–					
Nonacceptance	.42	–				
Impulse	.53	.46	–			
Clarity	.33	.46	.45	–		
Awareness	.05	.12	.15	.35	–	
Strategies	.58	.58	.65	.5	.1	–

All correlations were statistically significant at $p < .001$, with the exception of the correlation between the Goals and Awareness subscales, for which $p = .06$

adults) and .67 (sample 3, university adults). For samples 2 and 3, correlations between DERS-18 total score and SIDP total score ranged from .63 to .66. All correlations were statistically significant at $p < .001$.

To evaluate predictive validity of the DERS-18, we examine correlations of the original DERS and the DERS-18 total scores to positive and negative emotion regulation experiences during a subsequent 2-week daily diary study; these data were obtained only for sample 3 ($n = 113$). The original DERS was significantly correlated with difficulty regulating negative emotions ($r = .26, p = .006$) and positive emotions

($r = .21, p = .02$) during the diary period, and the DERS-18 exhibited near-identical correlations for both negative ($r = .27, p = .004$) and positive emotions ($r = .20, p = .03$).

Discussion

Understanding how individuals manage their emotional experiences is a critical part of affective science, and self-report measures of emotion dysregulation provide valuable information to researchers about these important constructs. One

Table 2 Factor loadings for the DERS-18 items using a six-factor model in the combined sample ($N = 1602$)

Original Item # (Subscale) ^a	Factors					
	1	2	3	4	5	6
2 I pay attention to how I feel. (A) (r)	.08	.02	.009	.10	.81	-.08
6 I am attentive to my feelings. (A) (r)	.04	-.06	-.03	-.006	.86	.10
10 When I am upset, I acknowledge my emotions. (A) (r)	-.12	.04	.02	-.08	.81	-.01
4 I have no idea how I am feeling. (C)	-.06	.05	.03	.86	.04	-.06
5 I have difficulty making sense out of my feelings. (C)	-.002	-.007	.06	.89	-.04	-.02
9 I am confused about how I feel. (C)	.03	-.03	-.08	.82	.00	.11
13 When I am upset, I have difficulty getting work done. (G)	.90	.01	.03	-.04	.02	.001
18 When I am upset, I have difficulty focusing on other things. (G)	.90	-.04	-.02	.01	-.03	.06
26 When I am upset, I have difficulty concentrating. (G)	.91	.05	.005	.00	.005	-.03
14 When I am upset, I become out of control. (I)	-.006	-.02	.90	.001	-.007	.03
27 When I am upset, I have difficulty controlling my behaviors. (I)	.13	.04	.87	.004	-.002	-.08
32 When I am upset, I lose control over my behaviors. (I)	-.083	-.01	.92	.006	.01	.08
12 When I am upset, I become embarrassed for feeling that way. (N)	.03	.87	-.03	.03	-.02	.02
21 When I am upset, I feel ashamed with myself for feeling that way. (N)	-.036	.90	.02	-.03	-.02	.06
25 When I am upset, I feel guilty for feeling that way. (N)	.03	.90	.009	.02	.03	-.04
15 When I am upset, I believe that I will remain that way for a long time. (S)	.01	-.06	.14	.01	-.03	.82
16 When I am upset, I believe that I'll end up feeling very depressed. (S)	.08	-.03	-.07	.05	-.008	.90
31 When I am upset, I believe that wallowing in it is all I can do. (S)	-.06	.17	.02	-.05	.05	.76

Bolded entries indicate factor loadings $> .75$

(r) refers to items that are reverse-coded for scoring purposes. For the subscales, *A* Awareness, *C* Clarity, *G* Goals, *I* Impulse, *N* Nonacceptance, *S* Strategies

^a These item numbers correspond to those used in the original, published form of the Difficulties in Emotion Regulation Scale (36-item). Note that these vary slightly from the item numbers reported in the original factor structure of the DERS by Gratz and Roemer (2004, p. 46), as psychometric data were reported for 41 items in the original paper, of which 36 were ultimately retained

Table 3 Reliability and concurrent validity of the DERS-18 subscale and overall scores across samples

	Samples											
	Combined sample		High school (Sample 1)		Adolescent inpatients (Sample 2)		University adults (Sample 3)		Community adults (Sample 4)		Community adults (Sample 5)	
	α	r	α	r	α	r	α	r	α	r	α	r
Awareness	.77	.92	.69	.9	.79	.92	.79	.92	.8	.91	.78	.93
Clarity	.83	.93	.78	.92	.8	.93	.81	.91	.81	.93	.87	.94
Goals	.9	.97	.89	.96	.86	.96	.92	.97	.89	.96	.91	.97
Impulse	.9	.95	.91	.95	.89	.94	.9	.96	.84	.93	.9	.95
Nonacceptance	.88	.95	.84	.91	.87	.95	.92	.94	.88	.96	.89	.97
Strategies	.85	.94	.8	.92	.82	.93	.84	.94	.84	.95	.87	.95
Overall	.91	.98	.87	.96	.9	.98	.9	.98	.91	.98	.92	.98

Values in the first column for each sample are the Cronbach's alpha (internal consistency reliability) for each DERS-SF subscale (or total score). Values in the second column for each sample are the Pearson's correlations between the DERS-SF subscale or total score and its corresponding original DERS subscale or total score (concurrent validity). All correlations are statistically significant at $p < .001$.

particular measure, the DERS (Gratz and Roemer 2004), is both well-validated and widely used, but requires significant participant time and effort to complete at its current length. We created a shortened form of the DERS and evaluated its reliability and validity to determine if the short version would measure emotion dysregulation in a shorter, simpler, and still reliable and valid way.

The revised DERS-18 appropriately reflected the six-factor structure of the original DERS in a large, combined sample as well as in several smaller samples that spanned age groups (adolescents, young adults, adults) and clinical status (non-clinical community participants, psychiatric patients). The DERS-18 exhibited high internal consistency, as well as strong convergent and concurrent validity by showing relationships with BPD symptoms and original DERS scores, respectively. The DERS-18 also demonstrates predictive validity insofar as greater difficulties with emotion regulation on the DERS-18 are associated with subsequent self-reported difficulties regulating both positive and negative emotions in daily life. These results suggest that the short form of the DERS is a reliable and valid measure of emotion dysregulation in a variety of populations.

This project has several strengths that improve confidence in our findings and ultimately support the use of the DERS-18. First, our factor-analytic methodology was selected to maintain consistency with the original DERS factor analysis. Second, we chose the strongest-loading items from the original DERS for use in the short-form, which helped maintain the six-factor structure and the high internal consistency reliability of each of the subscales. Third, we validated the full short form, as well as the short form's subscales, in multiple samples that differed not only in age range and clinical status, as previously noted, but also in collection methodology (in person versus online) and in racial/ethnic breakdown of the samples.

While we feel that the psychometric properties of the DERS-18 outlined here support its use by researchers interested in emotion dysregulation, there are some limitations that are worth noting. First, we could have selected items for inclusion in the DERS-18 based on each item's factor loading in our data, rather than based on the factor loadings reported by Gratz and Roemer (2004), which would have changed one item out of the 18 selected; however, this could have artificially inflated internal consistency estimates for the short form subscales and total score by capitalizing on unique or unusual responding patterns in our samples. Second, our results do not take into account the potential for fewer than six factors, as described by others who have evaluated the factor structure of the full DERS (e.g., Bardeen et al. 2012); while we hope that our decision to retain the original factors will provide greater consistency with existing work and will permit researchers to continue to evaluate multidimensional nature of emotion dysregulation, the existing factor structure may not, in fact, be the most parsimonious structure for the original DERS or for the DERS-18. Third, because our only clinical sample was comprised of adolescents, our ability to generalize our results to other clinical populations is somewhat limited. Fourth, we did not obtain test-retest reliability data for the DERS-18, so we cannot be entirely confident that the short form retains the test-retest reliability of the original form. Fifth, because all measures in the studies described here were related to emotions and psychopathology, there were no opportunities to investigate divergent validity with constructs that are unrelated to emotion dysregulation; establishing a broader range of convergent and divergent validity correlations would be a fruitful direction of future research.

In spite of these limitations, our results suggest that emotion dysregulation can be measured in a consistent, valid way with substantially fewer items than previously believed. All told, we believe that the shortened form of the DERS provides

researchers with a reliable and valid way to measure multiple dimensions of emotion dysregulation in a way that will facilitate the measure's inclusion in study protocols and reduce participant burden.

Compliance with Ethical Standards

Conflict of Interest Sarah E. Victor and E. David Klonsky declare that they have no conflicts of interest in the publication of this work.

Experiment Participants All participants provided informed consent (adults) or assent (minors) prior to participation in the study. For minors, parent/guardian consent was also obtained. All study procedures were conducted in accordance with appropriate institutional ethical review boards.

References

- Bardeen, J. R., & Fergus, T. A. (2014). An examination of the incremental contribution of emotion regulation difficulties to health anxiety beyond specific emotion regulation strategies. *Journal of Anxiety Disorders*. doi:10.1016/j.janxdis.2014.03.002.
- 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*. doi:10.1007/s10862-012-9280-y.
- Buhrmester, M., Kwang, T., & Gosling, D. (2011). Amazon's mechanical Turk: a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*. doi:10.1177/1745691610393980.
- Chapman, A. L., Leung, D. W., & Lynch, T. R. (2008). Impulsivity and emotion dysregulation in borderline personality disorder. *Journal of Personality Disorders*, 22(2). doi:10.1521/pepi.2008.22.2.148.
- Cho, Y., & Hong, S. (2013). The new factor structure of the Korean version of the difficulties in emotion regulation scale (K-DERS) incorporating method factor. *Measurement and Evaluation in Counseling and Development*. doi:10.1177/0748175613484033.
- Cooper, J. L., O'Shea, A. E., Atkinson, M. J., & Wade, T. D. (2014). Examination of the difficulties in emotion regulation scale and its relation to disordered eating in a young female sample. *International Journal of Eating Disorders*. doi:10.1002/eat.22278.
- Donahue, J. J., Goranson, A. C., McClure, K. S., & Van Male, L. M. (2014). Emotion dysregulation, negative affect, and aggression: a moderated, multiple mediator analysis. *Personality and Individual Differences*. doi:10.1016/j.paid.2014.06.009.
- 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*. doi:10.1016/j.jpsychires.2014.07.029.
- Franklin, J. C., Hessel, E. T., Aaron, R. V., Arthur, M. S., Heilbron, N., & Prinstein, M. J. (2010). The functions of nonsuicidal self-injury: support for cognitive-affective regulation and opponent processes from a novel psychophysiological paradigm. *Journal of Abnormal Psychology*. doi:10.1037/a0020896.
- Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*. doi:10.1093/poq/nfp031.
- Giromini, L., Velotti, P., de Campora, G., Bonalume, L., & Cesare Zavattini, G. (2012). Cultural adaptation of the difficulties in emotion regulation scale: reliability and validity of an Italian version. *Journal of Clinical Psychology*. doi:10.1002/jclp.21876.
- Glenn, C. R., & Klonsky, E. D. (2009). Emotion dysregulation as a core feature of borderline personality disorder. *Journal of Personality Disorders*. doi:10.1521/pepi.2009.23.1.20.
- Glenn, C. R., & Klonsky, E. D. (2013). Nonsuicidal self-injury disorder: an empirical investigation in adolescent psychiatric patients. *Journal of Clinical Child and Adolescent Psychology*. doi:10.1080/15374416.2013.79469.
- 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 Psychopathology and Behavioral Assessment*. doi:10.1023/B:JOBA.0000007455.08539.94.
- Gratz, K. L., Tull, M. T., & Levy, R. (2014). Randomized controlled trial and uncontrolled 9-month follow-up of an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality disorder. *Psychological Medicine*. doi:10.1017/S0033291713002134.
- Gross, J. J., & Jazaieri, H. (2014). Emotion, emotion regulation, and psychopathology: an affective science perspective. *Clinical Psychological Science*. doi:10.1177/2167702614536164.
- Joomann, J., & Vanderlind, M. (2014). Emotion regulation in depression: the role of biased cognition and reduced cognitive control. *Clinical Psychological Science*. doi:10.1177/2167702614536163.
- Kaufman, E. A., Xia, M., Fosco, G., Yaptangco, M., Skidmore, C. R., & Crowell, S. E. (2015). The Difficulties in Emotion Regulation Scale Short Form (DERS-SF): validation and replication in adolescent and adult samples. *Journal of Psychopathology and Behavioral Assessment*. doi:10.1007/s10862-015-9529-3.
- Klonsky, E. D., May, A. M., & Glenn, C. R. (2013). The relationship between nonsuicidal self-injury and attempted suicide: converging evidence from four samples. *Journal of Abnormal Psychology*. doi:10.1037/a0030278.
- Kököneyi, G., Urbán, R., Reinhardt, M., Józán, A., & Demetrovics, Z. (2014). The difficulties in emotion regulation scale: factor structure in chronic pain patients. *Journal of Clinical Psychology*. doi:10.1002/jclp.22036.
- 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*. doi:10.1016/j.janxdis.2009.04.006.
- Mitsopoulou, E., Kafetsios, K., Karademas, E., Papastefanakis, E., & Simos, P. G. (2013). The Greek version of the difficulties in emotion regulation scale: testing the factor structure, reliability and validity in an adult community sample. *Journal of Psychopathology and Behavioral Assessment*. doi:10.1007/s10862-012-9321-6.
- Pfohl, B., Blum, N., & Zimmerman, M. (1997). *Structured interview for DSM-IV personality*. Washington, DC: American Psychiatric Press.
- Racine, S. E., & Wildes, J. E. (2013). Emotion dysregulation and symptoms of anorexia nervosa: the unique roles of lack of emotional awareness and impulse control difficulties when upset. *International Journal of Eating Disorders*. doi:10.1002/eat.22145.
- Roszkowski, M. J., & Bean, A. G. (1990). Believe it or not! Longer questionnaires have lower response rates. *Journal of Business and Psychology*. doi:10.1007/BF01013611.
- Rottenberg, J., & Gross, J. J. (2003). When emotion goes wrong: realizing the promise of affective science. *Clinical Psychology: Science and Practice*. doi:10.1093/clipsy.bpg012.
- Ruganci, R. N., & Gençöz, T. (2010). Psychometric properties of a Turkish version of the difficulties in emotion regulation scale. *Journal of Clinical Psychology*. doi:10.1002/jclp.20665.
- Scott, L. N., Stepp, S. D., & Pilkonis, P. A. (2014). Prospective associations between features of borderline personality disorder, emotion

- dysregulation, and aggression. *Personality Disorders: Theory, Research, and Treatment*. doi:10.1037/per0000070.
- Sharp, C., Steinberg, L., Temple, J., & Newlin, E. (2014). An 11-item measure to assess borderline traits in adolescents: refinement of the BPFSC using IRT. *Personality Disorders: Theory, Research, and Treatment*. doi:10.1037/per0000057.
- Staples, A. M., & Mohlman, J. (2012). Psychometric properties of the GAD-Q-IV and DERS in older, community-dwelling GAD patients and controls. *Journal of Anxiety Disorders*, 26(3). doi:10.1016/j.janxdis.2012.01.005.
- Tracy, J. L., Klonsky, E. D., & Proudfit, G. H. (2014). How affective science can inform clinical science: an introduction to the special series on emotions and psychopathology. *Clinical Psychological Science*. doi:10.1177/2167702614537627.
- Victor, S. E., & Klonsky, E. D. (2014). Daily emotion in non-suicidal self-injury. *Journal of Clinical Psychology*. doi:10.1002/jclp.22037.
- Weinberg, A., & Klonsky, E. D. (2009). Measurement of emotion dysregulation in adolescents. *Psychological Assessment*. doi:10.1037/a0016669.
- Zanarini, M. C., Vujanovic, A. A., Parachini, E. A., Boulanger, J. L., Frankenburg, F. R., & Hennen, J. (2003). A screening measure for BPD: the mclean screening instrument for borderline personality disorder (MSI-BPD). *Journal of Personality Disorders*. doi:10.1521/pedi.17.6.568.2535z.