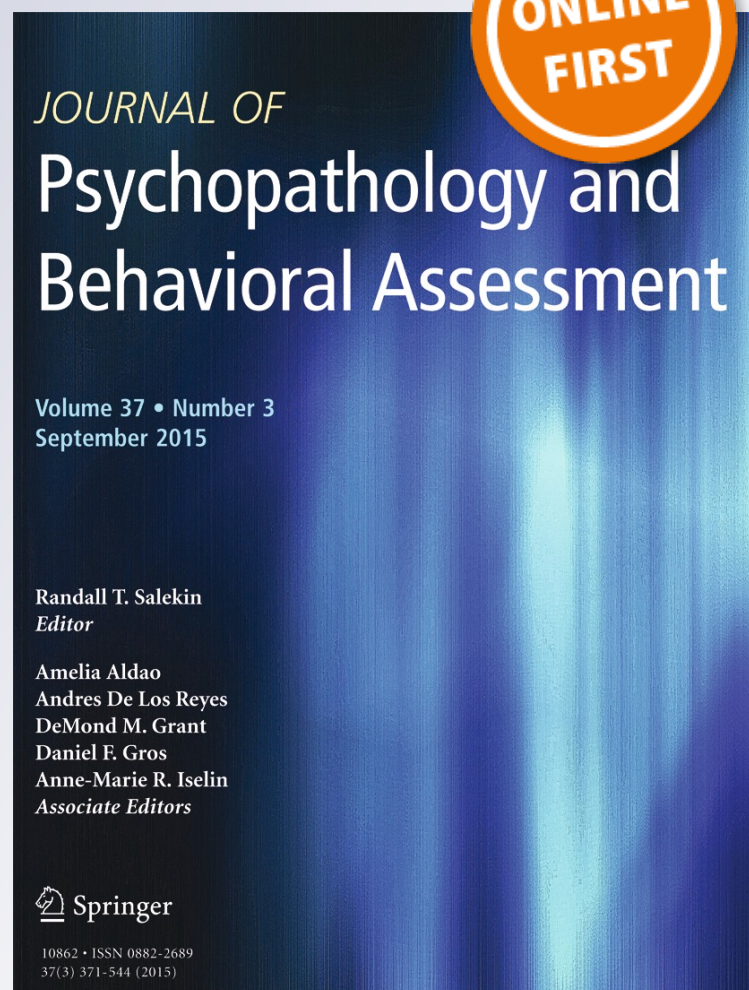# The Difficulties in Emotion Regulation Scale Short Form (DERS-SF): Validation and Replication in Adolescent and Adult Samples

Erin A. Kaufman, Mengya Xia, Gregory Fosco, Mona Yaptangco, Chloe R. Skidmore & Sheila E. Crowell

JOURNAL OF

Psychopathology and
Behavioral Assessment

Volume 37 • Number 3
September 2015

Randall T. Salekin
Editor

Amelia Aldao
Andres De Los Reyes
DeMond M. Grant
Daniel F. Gros
Anne-Marie R. Iselin
Associate Editors

Springer

10862 • ISSN 0882-2689
37(3) 371-544 (2015)

Springer

Springer

CrossMark

The Difficulties in Emotion Regulation Scale Short Form (DERS-SF): Validation and Replication in Adolescent and Adult Samples

Erin A. Kaufman[1] · Mengya Xia[2] · Gregory Fosco[2] · Mona Yaptangco[1] ·
Chloe R. Skidmore[1] · Sheila E. Crowell[1,3]

Abstract Emotion dysregulation often emerges early in development and is a core feature of many psychological conditions. The Difficulties in Emotion Regulation Scale (DERS) is a well validated and widely used self-report measure for assessing emotion regulation problems among adolescents and adults. The DERS has six subscales with five to eight items each (36 total), suggesting multiple questions may assess similar underlying constructs. In an effort to reduce respondent burden and streamline this widely-used instrument, we created a short-form version of the DERS (DERS-SF) using confirmatory factor analysis on data from three adolescent ($n=257$) and two adult samples ($n=797$). Scores on the DERS-SF yielded similar correlation patterns relative to the full measure, ranging from .90 to .98 and reflecting 81–96 % shared variance. This instrument maintains the excellent psychometric properties and retains the total and subscale scores of the original measure with half the items.

Keywords Emotion regulation · Assessment · Factor analysis · Short-form · Validation

The ability to effectively regulate emotions is a core competency for healthy functioning (Cole 2014). When emotion regulation skills are underdeveloped or otherwise compromised, normative affective development may be delayed and risk for psychopathology increases (McLaughlin et al. 2011). Broadly defined, emotion regulation includes the ability to: identify, understand, and accept emotional experiences, control impulsive behaviors when distressed, and flexibly modulate emotional responses as situationally appropriate (Cole et al. 1994; Gratz and Roemer 2004; Eisenberg and Spinrad 2004; Linehan 1993; Thompson 1994). These abilities typically increase with age (Orgeta 2009). However, difficulties with emotion regulation occur across the lifespan.

Emotion dysregulation is a core feature of disorders that span the internalizing and externalizing spectra (Beauchaine and Thayer 2015; Hofmann et al. 2012). Researchers have observed links between emotion dysregulation and self-inflicted injury (Gratz and Tull 2010; Crowell et al. 2005), identity disturbance (Kaufman et al. 2015), substance abuse (e.g., Dvorak et al. 2014), depression (Crowell, et al. 2014), conduct problems (Beauchaine et al. 2007; Cappadocia et al. 2009), attention-deficit/hyperactivity disorder (Mitchell et al. 2012), anxiety (Folk et al. 2014), post-traumatic stress (Weiss et al. 2013), borderline personality disorder (Fossati et al. 2014), and eating disorders (Lavender et al. 2014; Racine and Wildes 2013). Thus, emotion dysregulation is an excellent transdiagnostic indicator of vulnerability and may contribute to high rates of comorbidity across various diagnoses (Beauchaine and Thayer 2015).

The Difficulties in Emotion Regulation Scale

The Difficulties in Emotion Regulation Scale (DERS; Gratz and Roemer 2004) is one of the most widely used self-report measures of emotion regulation deficits. The DERS was developed to capture clinically relevant problems (i.e., those

✉ Sheila E. Crowell
sheila.crowell@psych.utah.edu

[1] Department of Psychology, University of Utah, 380 S. 1530 E. BEHS 502, Salt Lake City, UT 84112, USA

[2] Human Development and Family Studies, The Pennsylvania State University, State College, PA, USA

[3] Department of Psychiatry, University of Utah, Salt Lake City, UT, USA

Springer

significant enough to be associated with clinical diagnosis; Gratz and Roemer 2004). However, it has also been used to examine normative developmental processes and experiences. For example, identity development, procrastination, social participation, academic motivation and performance, and psychophysiological responding are each associated with emotion regulation functioning as assessed by the DERS (Jankowski 2013; Jankowski and Rękosiewicz 2013; Mirzaei et al. 2014; Singh and Singh 2013). The measure has good reliability and validity with adolescents (Neumann et al. 2010; Sarıtaş-Atalar et al. 2015; Weinberg and Klonsky 2009) and adults of all ages (Kökönyei et al. 2014; Orgeta 2009; Staples and Mohlman 2012). It has also been validated with international samples and translated into multiple languages (e.g., Côté et al. 2013; Fossati et al. 2014; Sarıtaş-Atalar et al. 2014).

The DERS consists of 36 items that load onto six subscales (Gratz and Roemer 2004). *Nonacceptance* of emotional responses reflects a tendency toward negative secondary responses to negative emotions, and/or denial of distress. The *difficulties engaging in goal-directed* behavior scale captures problems concentrating and accomplishing tasks while experiencing negative emotions. The *impulse control difficulties* subscale reflects struggles to control behavior when upset. The *lack of emotional awareness* scale captures inattention to emotional responses. The *limited access to emotion regulation strategies* scale assesses beliefs that there is little a person can do to regulate one's emotions effectively after becoming upset. Finally, the last subscale, labeled *lack of emotional clarity*, reflects the extent to which individuals are unclear about which emotions they are experiencing.

Although the DERS is a useful and widely-studied instrument, many of the items are conceptually similar. DERS subscales contain between five and eight statements that load strongly on to each subscale, suggesting that multiple items may not be necessary to adequately assess the underlying constructs. Furthermore, the similarity of some items may be perceived as repetitive to participants, potentially increasing frustration and fatigue. The nuances captured by each question in the DERS may be useful for research focused explicitly on emotion dysregulation. Nonetheless, a shortened version of this instrument could effectively capture core dimensions of emotion dysregulation. Meta-analytic findings indicate participant response rates are generally lower for lengthier questionnaires (Edwards, et al. 2002; Rolstad et al. 2011). In addition, brief measures can be equally or more valid than longer measures of the same construct (Smith et al. 2012). Given the broad significance of emotion regulation difficulties, measuring this construct efficiently may prove useful in epidemiological studies.

The primary goal of this study was to develop and validate a short form of the DERS (the DERS-SF). Our aim was to preserve the factor structure, achieve comparable psychometric properties and concurrent validity of the full version, and reduce survey length by 50 %. This would allow for less item redundancy, simplified scoring, shorter assessment times, and less respondent burden. In the studies presented here, we utilized data from five independent samples (presented as two CFA studies) evaluating the utility of the DERS-SF in adolescents and adults. Specifically, we used confirmatory factor analysis to examine DERS-SF items with three pooled adolescent samples (Study 1) and then sought to replicate our findings with two pooled college samples (Study 2).

We tested the following hypotheses:

H1: *It would be possible to create a psychometrically sound DERS-SF using the same 6-factor structure as the original measure.* Consistent with this hypothesis, we expect the CFA for a short form of the DERS will have similar to improved model fit compared with a CFA of the original DERS. We predict items in the DERS-SF will have similar factor loadings across DERS and DERS-SF CFA's, and expect the DERS-SF will account for a large proportion of variance in emotion dysregulation explained by the original DERS. Accordingly, we expect that correlations between DERS and DERS-SF subscales will be high (e.g., r>.90), and predict the within-scale correlations among subscales will be similar across the two versions of the measure.

H2: *The DERS-SF and the original DERS will offer comparable validity as transdiagnostic indicators of risk for psychopathology.* Consistent with this hypothesis, we compute correlations of DERS and DERS-SF subscales with established measures of psychopathology, allowing us to compare concurrent validity for the two DERS forms in relation to psychopathology risk. We expect correlations between DERS-SF subscales and key psychopathology outcomes will be comparable to the correlations between DERS subscales with the same outcome measures.

H3: *The DERS-SF will perform similarly across adolescent and college samples.* Consistent with this hypothesis, we conduct parallel tests of the DERS-SF in adolescent and college samples, and expect the DERS-SF to offer comparable psychometric properties and validity across developmental periods.

Method

Overview of Studies

Data were collected from five independent samples in the Pacific Northwest and Mountain West regions of the United States. Three are adolescent samples, which we combined and reported as *Confirmatory Factor Analysis (CFA) Study 1*. CFA

Study 2 included two samples to replicate and extend our findings into adulthood. Institutional review board approval and informed assent or consent were obtained for all studies and participants.

CFA Study 1: Initial Validation of the DERS-SF Among Adolescents

Sample 1

Participants This sample and consisted of 84 adolescent girls placed into one of three groups: self-injuring ($n=29$), depressed with no self-injury history ($n=28$), and typical control ($n=27$). Participants were age 13–18 ($M=16.04$, $SD=1.24$), with 71.1 % identifying as Caucasian, 3.6 % as African American, 4.8 % as Hispanic, 6 % as Asian, 10.8 % as mixed racial/ ethnic heritage, 1.2 % as other, and 2.4 % declining to answer. Participants for this sample were taken from a larger, multi-visit study that included the adolescents' mothers. Although mothers' self-report data were not used for the current study, their informant ratings of adolescent functioning were collected (see *Measures*). Participants were recruited using mailers, classified ads, banners on busses, and brochures distributed to local schools, inpatient treatment facilities, and outpatient clinics. Participants were compensated $60 for the full study.

Sample 2

Participants Participants included 131 adolescents ages 13 to 18 ($M=15.3$, $SD=1.18$), recruited for an online survey. Due to incomplete survey data, 17 participants were excluded, leaving a final sample of 114. The sample was 63.2 % female, 81.6 % Caucasian, 6.1 % Asian, 6.1 % Hispanic, and 6.2 % Other (Pacific Islander, African American, or mixed racial/ ethnic heritage). Participants were recruited via mailers addressed to parents/guardians that included an invitation to participate, a consent document, and an address for the survey web-link. Adolescents who were interested in participating were directed to a fully online informed assent document, where they acknowledged their assent and were assigned an individual identification code. Participants were then linked to a separate secure, de-identified survey. As compensation, adolescents were entered in a drawing for each page of the survey they completed, for a total of seven possible entries.

Sample 3

Participants Sample 3 was taken from a family study comparing adolescent suicide attempters ($n=29$) with typical controls ($n=30$). Adolescents ranged in age from 12 to 20 and were 70.4 % female, 91.5 % Caucasian, 1.7 % African American, and 6.8 % of mixed racial/ethnic heritage. Six adolescents (10.2 %) identified as Hispanic or Latino, with the remaining 89.8 % identifying as non-Hispanic. Suicidal adolescents were recruited from outpatient or inpatient psychiatry at a local neuropsychiatric hospital, adolescent medicine, and online classifieds (e.g., Craigslist). Controls were recruited through flyers on community websites and classifieds, local businesses, community centers, organizations (e.g., Boy Scouts), libraries, pediatric clinics, mailings, and word of mouth. Parent reports of adolescent functioning were collected (see *Measures*). Participants were compensated up to $50 for the full study.

Measures All adolescents completed *the Difficulties in Emotion Regulation Scale.* The DERS consists of 36 items that load onto 6 subscales. In order to assess difficulties regulating emotions during times of distress, many items begin with "When I'm upset." Respondents are asked to indicate how often the items apply to themselves, 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 consistency ($\alpha=.93$), good test-retest reliability ($\rho=.88$, $p<.01$), and adequate construct and predictive validity (Gratz and Roemer 2004). The internal reliability for the current sample was .95.

Each sample also completed a number of other outcome measures, which we used to test the concurrent validity of the DERS-SF relative to the full DERS. Selected measures assess constructs that are theoretically and empirically linked to emotion dysregulation, and scores on many of these instruments have previously shown high correlations with the scores on the DERS.

Adolescents in samples 1 and 3 provided ratings of their behavior problems, psychopathology and broadband internalizing and externalizing symptoms on the *Youth Self-Report* (Achenbach 1987). The YSR is 112-items rated on a three-point scale (0=never, 1=sometimes, 2=often). The YSR is widely used and is well-validated, with excellent psychometric properties (Achenbach 1991a). Others have reported internal consistency reliabilities (Cronbach α) that range from .72 to .93 (Gadow et al. 2002). Estimates in the current sample ranged from a low of .76 for the social problems subscale to .89 for the anxious/depressed subscale.

Parent ratings of child behavior problems were also available for Samples 1 and 3 via the *Child Behavior Checklist (CBCL;* the parent-report version of the YSR, Achenbach 1991b). The CBCL is among the most widely used measures of child psychopathology. The 113 items of the CBCL were factor analyzed to empirically identify forms of adolescent psychopathology, including broad internalizing and externalizing scales. Prior research has yielded acceptable construct validity for internalizing ($r=.56$ to .72) and externalizing behaviors ($r=.52$ to .88; Achenbach 1991b), and acceptable to high internal consistency across all behavioral scales for a matched sample of referred and non-referred 12–18 year-old

Table 1 Descriptive statistics for adolescent sample

Adolescent	N		M	SD	Range	
DERS	Strategies	257	2.28	1.08	1.00–5.00	
	Nonacceptance	257	2.16	1.10	1.00–5.00	
	Impulse	257	2.02	0.84	1.00–4.83	
	Goals	257	2.82	1.13	1.00–5.00	
	Awareness	257	2.84	1.00	1.00–5.00	
	Clarity	257	2.48	0.96	1.00–5.00	
	Total	257	2.41	0.77	1.00–4.39	
DERS-SF	Strategies	257	2.17	1.17	1.00–5.00	
	Nonacceptance	257	2.08	1.14	1.00–5.00	
	Impulse	257	1.62	0.92	1.00–5.00	
	Goals	257	2.80	1.25	1.00–5.00	
	Awareness	257	2.57	1.10	1.00–5.00	
	Clarity	257	2.18	1.08	1.00–5.00	
	Total	257	2.24	0.79	1.00–4.67	
YSR	Withdrawn	257	60.08	9.84	50–96	
	Somatic complaints	257	56.63	8.31	50–90	
	Anxious/depressed	257	61.21	11.38	50–95	
	Social problems	257	59.89	8.82	50–85	
	Thought	257	59.72	9.16	50–95	
	Attention	257	59.03	9.50	50–91	
	Rule breaking	257	59.34	8.46	50–91	
	Aggressive behavior	257	56.08	7.47	50–82	
	Externalizing scale	257	55.59	10.76	29–83	
	Internalizing scale	257	57.63	13.30	27–91	
CBCL	Withdrawn	140	61.04	10.52	50–96	
	Somatic complaints	137	60.09	9.46	50–92	
	Anxious/depressed	139	60.35	10.80	50–90	
	Social problems	137	55.24	6.91	50–78	
	Thought	137	57.40	8.51	50–80	
	Attention	137	57.33	8.58	50–86	
	Rule breaking	137	58.17	9.48	50–90	
	Aggressive behavior	135	56.22	7.96	50–93	
	Externalizing scale	133	52.53	12.91	33–82	
	Internalizing scale	135	58.67	13.79	31–83	
SCIM	Total	140	2.96	1.00	1.26–5.19	
	Consolidated	144	2.77	0.87	1.13–5.20	
	Disturbed	144	3.12	1.11	1.00–6.00	
	Lack of Self	144	2.86	1.71	1.00–7.00	
Self-Harm		252	0.35	0.48	0–1	

DERS, DERS-SF, SCIM and Self-Harm data are the means of average score; YSR, CBCL are the means of total score

children (α=.68–.92; Achenbach 1991b). Test–retest reliability coefficients were also high ranging from .79 to .92 (Achenbach, Edelbrock, and Howell 1987). The internal reliability for the current sample was excellent (Cronbach's α=.92 for internalizing and externalizing scales in sample 1; Cronbach's α=.94 and .95 for the internalizing and externalizing scales respectively in sample 3). As stated above,

emotion dysregulation is a key factor underlying psychopathology. Thus the YSR and CBCL were included to assess the concurrent validity of the DERS-SF. Previous research has demonstrated that the DERS correlates strongly with scores on these measures (see e.g., Vasilev, Crowell, Beauchaine, Mead, and Gatzke-Kopp 2009), therefore CBCL/YSR scores were also expected to correlate strongly with DERS-SF scores.

Participants from Samples 2 and 3 completed the *Self-Concept and Identity Measure (SCIM;* Kaufman et al. 2015). The SCIM assesses the core aspects of identity such as: self-concept and role continuity across environments and persons,

Table 2 Descriptive statistics for college sample

College	N		M	SD	Range	
DERS	Strategies	794	1.95	0.80	1.00–5.00	
	Nonacceptance	796	2.00	0.87	1.00–5.00	
	Impulse	794	1.67	0.79	1.00–5.00	
	Goals	795	2.68	0.96	1.00–5.00	
	Awareness	796	2.27	0.80	1.00–4.67	
	Clarity	796	2.01	0.68	1.00–5.00	
	Total	791	2.08	0.62	1.00–4.32	
DERS-SF	Strategies	794	1.89	0.87	1.00–5.00	
	Nonacceptance	796	1.97	0.90	1.00–5.00	
	Impulse	795	1.54	0.78	1.00–5.00	
	Goals	796	2.73	1.06	1.00–5.00	
	Awareness	796	2.07	0.84	1.00–4.67	
	Clarity	796	1.82	0.70	1.00–5.00	
	Total	791	2.00	0.61	1.00–4.59	
BDI-II		794	9.90	8.82	0–47	
SCL-90-R	Obsessive Compulsive	792	0.52	0.64	0.00–3.17	
	Interpersonal Sensitivity	791	0.53	0.63	0.00–3.75	
	Depression	791	0.33	0.44	0.00–3.00	
	Anxiety	791	0.59	0.63	0.00–3.33	
	Hostility	791	0.63	0.67	0.00–3.40	
	Phobic Anxiety	791	0.52	0.59	0.00–3.40	
	Paranoid Ideation	791	0.47	0.61	0.00–3.60	
	Psychoticism	791	0.68	0.78	0.00–3.80	
	Somatization	722	0.60	0.72	0.00–3.71	
	GSI	791	0.55	0.57	0.00–3.08	
STAI-Y	State	229	33.77	9.89	20–67	
	Trait	229	35.55	10.57	20–72	
AAQ-II		225	54.62	9.76	20–70	
SCIM	Total	535	68.65	21.15	31–136	
	Consolidated	535	24.83	7.26	12–65	
	Disturbed	567	29.71	11.03	11–77	
	Lack of Self	567	14.22	8.27	6–44	
Self-Harm		793	0.28	0.45	0–1	

DERS, DERS-SF, SCL-90-R, and Self-Harm data are the means of average score; BDI-II, STAI-Y, AAQ-II, and SCIM are the means of total score

consistencies in values and interests, self-worth, self-differen-tiation, and self-cohesion. It assesses adaptive and maladap-tive identity functioning across three subscales. Internal con-sistency of the scale is excellent (Cronbach's α=.89). Test-retest reliability is also excellent (α=.93, r=.88) with an intraclass correlation coefficient (ICC) of .88. Internal reliabil-ity for the total scale in our adolescent sample was acceptable at .69, and the disturbed identity, consolidated identity, and lack of identity subscales were good to excellent (α=.83, .85, and .93 respectively). Identity problems have been theo-retically and empirically linked to emotion dysregulation, and previous research has demonstrated moderate to strong corre-lations between SCIM and original DERS scores (Kaufman et al. 2015). We anticipated DERS-SF scores would yield similar correlation patterns with this outcome measure.

Finally, self-injury history was obtained via a single item on the *Lifetime Suicide Attempt Self-Injury Interview* (L-SASII; formerly the Lifetime Parasuicide Count; Linehan and Comtois 1996) for samples 1 and 3. The same item was used for sample 2 but was modified slightly (i.e., revised for questionnaire-based responding rather than an interview). The specific item ("Have you ever intentionally injured yourself?") is scored dichotomously (i.e., positive or negative for self-harm history). This item was included as an additional measure of concurrent validity given that emotion dysregula-tion is a key risk-factor for self-inflicted injury (see e.g., Crowell et al. 2005; Gratz and Tull 2010). The DERS-SF was expected to be positively associated with history of self-harm behavior.

Procedure All adolescents completed questionnaires and de-mographic questions either during a laboratory visit (Samples 1 and 3), or on a personal computer at home (Sample 2).

CFA Study 2: Replication Among College Students

Sample 1

Participants Participants for Sample 1 were 230 students en-rolled in undergraduate psychology courses at a large univer-sity. The participants ranged in age from 19 to 64 years (M= 24.38, SD=5.80). Approximately 63 % of the sample was female, and 88.7 % of the sample identified as Caucasian, 7.8 % Asian, .9 % African American, and .9 % as American

Table 3 Confirmatory factor loadings for the DERS and DERS-SF in adolescent and college samples

	Adolescent (n= 257)		College (n= 797)	
	Full	SF	Full	SF
Strategies				
35 When I'm upset, it takes me a long time to feel better.	.82	.80	.76	.75
28 When I'm upset, I believe there is nothing I can do to make myself feel better.	.84	.86	.79	.81
16 When I'm upset, I believe that I will end up feeling very depressed.	.84	.82	.85	.88
Non-acceptance				
12 When I'm upset, I become embarrassed for feeling that way.	.79	.77	.77	.74
25 When I'm upset, I feel guilty for feeling that way.	.81	.79	.85	.82
29 When I'm upset, I become irritated at myself for feeling that way.	.89	.92	.83	.86
Impulse				
14 When I'm upset, I become out of control.	.81	.81	.84	.81
32 When I'm upset, I lose control over my behavior.	.86	.88	.85	.87
27 When I'm upset, I have difficulty controlling my behavior.	.84	.85	.84	.88
Goals				
18 When I'm upset, I have difficulty focusing on other things.	.87	.87	.88	.90
26 When I'm upset, I have difficulty concentrating.	.86	.88	.88	.88
13 When I'm upset, I have difficulty getting work done.	.84	.84	.84	.84
Awareness				
8 I care about what I am feelings.	.72	.75	.81	.84
10 When I'm upset, I acknowledge my emotions.	.68	.73	.64	.63
2 I pay attention to how I feel.	.83	.76	.81	.76
Clarity				
9 I am confused about how I feel.	.73	.73	.61	.68
5 I have difficulty making sense out of my feelings.	.87	.93	.72	.84
4 I have no idea how I am feeling.	.82	.80	.66	.76

Indian or Alaska native. Four participants did not report their race. Participants were recruited through a department participant pool and received $30 and research credit for their time. Given the large number of non-traditional students at the university, there was no upper limit placed on age.

Sample 2

Participants Participants for Sample 2 were also students enrolled in undergraduate psychology courses at the same university. The final sample of 567 participants was 18 to 65 years of age (M=24.20; SD=6.21; 43 % male). Demographics were missing for 186 participants (32.8 % of the sample) due to an error in the survey. Of the participants with demographics, 79.3 % identified as Caucasian, 6.8 % as Asian, .8 % as Pacific Islander, 1.3 % as African American, 5.2 % as Hispanic, 5.8 % as multiple races, and .8 % reported other/unspecified race.

Measures In addition to the DERS (Cronbach α=.94 for the current sample) and the SCIM (Cronbach α=.89 for Study 2 Sample 2; the SCIM was not collected in Sample 1), participants for this study also completed the *Beck Depression Inventory-II (BDI-II)*. The BDI-II is a 21-item questionnaire used to measure depression severity (Beck et al. 1996). Because emotion dysregulation is implicated in risk for depression, we used this instrument as another important index of concurrent validity for the DERS-SF. Scores on the BDI-II have excellent psychometric properties, and have demonstrated high test-retest reliability (r=0.93, p<0.001; Beck et al. 1996). Internal reliability within the current sample was excellent (α=.91).

Self-injury history in Samples 1 and 2 was obtained via *the Deliberate Self-Harm Inventory (DSHI),* a self-report questionnaire of non-suicidal self-harm behavior used in the original DERS validation study (Gratz 2001; Gratz and Roemer 2004). For this study, participant responses were coded using 17 dichotomous items (e.g., "Have you ever intentionally [i.e., on purpose] cut your wrist, arms, or other area (s) of your body?"). Endorsement of any self-harm method was coded positive for self-harm history and declining all methods resulted in a negative code. We opted to reduce the different methods of self-harm into a single yes/no variable in order to be consistent with the approach taken in the adolescent sample. We expected that DERS-SF scores would be positively associated with reported history of self-harm behavior.

Participants completed *the Symptom Checklist-90-Revised (SCL-90-R),* which measures nine symptom dimensions, their severity, and provides a Global Severity Index (GSI) as a measure of overall psychopathology (Derogatis and Lazarus 1994). Previous research has demonstrated that scores on the SCL-90-R have good to excellent internal consistency (Cronbach's α ranging from .69 to .91; Buckelew, Burk,

Brownlee-Duffeck, Frank & DeGood, 1988; Derogatis and Lazarus 1994; Kaufman et al. 2015), and high test-retest reliability (Cronbach's α ranging from .78 to .90; Derogatis & Lazarus). Cronbach's α in the current sample ranged from acceptable to excellent (.70 for the hostility scale to .91 for the depression scale). As with the CBCL and YSR in our adolescent sample, psychopathology as measured by the SCL-90 was expected to correlate positively with DERS-SF scores.

Participants from Sample 2 also completed the *Acceptance and Action Questionnaire- II (AAQ- II;* Hayes et al. 2004) and the *State and Trait Anxiety Inventory-Form Y (STAI-Y;*

Table 4 Within-measure subscale correlations for the DERS and DERS-SF in adolescent and college samples

Adolescent (n= 257): Correlation Between Scales in DERS-SF and DERS						
DERS-SF Correlations	1	2	3	4	5	6
Strategies	–					
Nonacceptance	.58**	–				
Impulse	.56**	.49**	–			
Goals	.61**	.43**	.47**	–		
Awareness	.26**	.20**	.17**	.04	–	
Clarity	.57**	.52**	.39**	.45**	.31**	–
Total	.85**	.76**	.70**	.72**	.46**	.76**
DERS Correlations	1	2	3	4	5	6
Strategies	–					
Nonacceptance	.66**	–				
Impulse	.69**	.54**	–			
Goals	.67**	.45**	.58**	–		
Awareness	.32**	.24**	.25**	.14*	–	
Clarity	.60**	.54**	.46**	.45**	.51**	–
Total	.90**	.78**	.78**	.73**	.53**	.77**
College (n= 797): Correlations Between Scales in DERS-SF and DERS						
DERS-SF Correlations	1	2	3	4	5	6
Strategies	–					
Nonacceptance	.57**	–				
Impulse	.66**	.53**	–			
Goals	.61**	.44**	.49**	–		
Awareness	.26**	.21**	.26**	.11**	–	
Clarity	.54**	.54**	.48**	.40**	.37**	–
Total	.82**	.76**	.77**	.73**	.51**	.75**
DERS Correlations	1	2	3	4	5	6
Strategies	–					
Nonacceptance	.68**	–				
Impulse	.76**	.57**	–			
Goals	.68**	.46**	.55**	–		
Awareness	.31**	.23**	.25**	.15**	–	
Clarity	.60**	.51**	.52**	.44**	.59**	–
Total	.90**	.77**	.80**	.73**	.55**	.78**

*p<.05, **p<.01

Spielberger et al. 1980) as additional measures of the concurrent validity of the DERS-SF and its similarity to the original measure. The AAQ-II assesses experiential avoidance (i.e., the urge to avoid distressing thoughts, feelings, and sensations) and has sufficient test-retest reliability ranging from .79 to .81, good discriminant validity (Bond et al. 2011), and high internal reliability within the current sample (Cronbach's α=.90). Avoidance is a frequently used maladaptive emotion regulation strategy, and this measure was used in Gratz and Roemer's original validation study for the DERS (2004). The AAQ was expected to be positively associated with scores on this instrument. The STAI-Y is a 40-item measure of both state and trait anxiety with excellent reliability (α=.89) and test-retest reliability (ρ=.88; Barnes, Harp, and Jung 2002). The STAI-Y was included to assess concurrent validity and was expected to correlate with the DERS-SF. Internal reliability within the current sample was excellent (α=.93 for state anxiety and .92 for trait anxiety).

Procedure Participants for both samples completed self-report measures through a secure online survey hosted by the University's psychology department. After completing these questionnaires, participants were debriefed, compensated monetarily and/or with credit hours, and provided with a page long list of mental health resources throughout the community.

DERS-SF Development

In order to identify which items best represent the original DERS subscales, we first examined published reports by

Gratz and Roemer (2004; the original validation study), Neumann, van Lier, Gratz, and Koot (2010; a replication study with adolescents ages 11–17) and Weinberg and Klonsky (2009; a replication study with adolescents ages 13–17). Both studies reported item-level exploratory factor analyses (EFA) of the DERS in their respective populations. The 6-factor structure was upheld in both samples, with acceptable subscale reliability. However, there was some variability in the strength of the factor loadings for the items within the scales for adult and adolescent samples. Thus, we were interested in identifying items that would perform well for both adult and adolescent populations in our shortened scales.

Item selection for the confirmatory factor analysis (CFA) was conducted by examining published EFA results for adolescent (Neumann, et al. 2010; Weinberg and Klonsky 2009) and college student samples (Gratz and Roemer 2004). We used three empirical approaches to identifying items that would perform well for adolescents and adults, have a strong factor loading on the primary scale, and have minimal cross-loading on other scales. Three metrics were used to provide breadth of empirical indicators for strongly loading items. First, we computed an average rank order score within each subscale. This was conducted by ranking items from the largest factor loading to the smallest for each study. We then averaged these values in the adult and adolescent studies. This approach gave preference to the "best items" in each scale. Second, we computed an average factor loading for the two studies. Third, we computed an average "discrimination score" to compare the difference between how well an item loads on its primary factor and it's loading on "off" factors (see Donnellan, Oswald, Baird, and Lucas 2006 for another

Table 5 Correlation between DERS-SF and DERS, reliability, and descriptive statistics for adolescent and college samples

Adolescent (*n*= 257)	Correlation between DERS and DERS-SF	DERS			DERS-SF		
		α	Mean	SD	α	Mean	SD
Strategies	.96	.92	2.28	1.08	.87	2.17	1.17
Nonacceptance	.97	.92	2.16	1.10	.87	2.08	1.14
Impulse	.91	.82	2.02	0.84	.88	1.62	0.92
Goals	.96	.89	2.82	1.13	.90	2.80	1.25
Awareness	.94	.84	2.84	1.00	.79	2.57	1.10
Clarity	.93	.85	2.48	0.96	.86	2.18	1.08
Total Score	.98	.95	2.41	0.77	.91	2.24	0.79
College (*n*= 797)	Correlation between DERS and DERS-SF	DERS			DERS-SF		
		α	Mean	SD	α	Mean	SD
Strategies	.91	.89	1.95	.80	.82	1.88	.87
Nonacceptance	.96	.91	2.00	.87	.85	1.97	.90
Impulse	.95	.86	1.67	.79	.89	1.54	.78
Goals	.96	.89	2.68	.96	.91	2.73	1.06
Awareness	.93	.85	2.27	.80	.78	2.07	.84
Clarity	.90	.81	2.01	.68	.78	1.82	.70
Total Score	.97	.94	2.08	.62	.89	2.00	.61

application of this approach). An average of the absolute value of "off" factor loadings is computed and subtracted from the primary factor loading. As an example, the item "When I'm upset, I feel guilty for feeling that way" loads primarily on the Nonacceptance scale. The factor loading reported by Gratz and Roemer (2004) was .91. The average of the absolute values of the loadings on the other 5 scales was .11. Thus, the discrimination score for this item was .80. Finally, when multiple items were identified as equally strong candidates using the three empirical approaches, we gave preference to breadth in topical coverage to reduce repetition of item content.

The three empirical metrics arrived at a clear consensus for the three preferred items in 4 of 6 scales: the Strategies, Impulse, Goals, and Clarity scales. In the Non-acceptance scale, the top three items were related to feeling embarrassed,

ashamed, and guilty for becoming upset. To allow for a greater breadth of conceptual coverage, we opted to include "irritated" and drop "ashamed" for the CFA analyses. Finally, for the Awareness scale, both "I am attentive to my feelings" and "I pay attention to how I feel" were strong items, but deemed repetitive. Thus, "I am attentive" was excluded in favor of including "I pay attention to how I feel."

Results

Study 1: Validation of the DERS-SF With an Adolescent Sample

Descriptive statistics for all self-report measures in Study 1 are summarized in Table 1.

Table 6 Comparisons of concurrent validity for DERS and DERS-SF for adolescent sample

		Strategies		Non-Accept.		Impulse		Goals		Awareness		Clarity		Total	
		Full	SF	Full	SF	Full	SF	Full	SF	Full	SF	Full	SF	Full	SF
YSR[a]	Withdrawn	.69**	.69**	.57**	.50**	.45**	.35**	.45**	.43**	.33**	.31**	.57**	.56**	.69**	.67**
	Somatic complaints	.62**	.60**	.48**	.45**	.48**	.41**	.43**	.39**	.26**	.24**	.44**	.46**	.61**	.60**
	Anxious/ depressed	.77**	.74**	.69**	.65**	.61**	.52**	.52**	.48**	.30**	.26**	.56**	.58**	.78**	.76**
	Social problems	.61**	.60**	.53**	.50**	.49**	.42**	.41**	.39**	.30**	.28**	.54**	.54**	.65**	.64**
	Thought	.61**	.60**	.55**	.51**	.52**	.45**	.38**	.37**	.28**	.27**	.53**	.54**	.64**	.64**
	Attention	.50**	.49**	.46**	.42**	.49**	.40**	.48**	.46**	.27**	.24**	.47**	.47**	.59**	.59**
	Rule breaking	.53**	.49**	.40**	.36**	.48**	.45**	.37**	.34**	.18**	.18**	.42**	.43**	.53**	.53**
	Aggressive behavior	.51**	.46**	.42**	.38**	.58**	.54**	.37**	.34**	.18**	.17**	.36**	.37**	.54**	.53**
	Externalizing scale	.58**	.54**	.47**	.42**	.57**	.50**	.45**	.43**	.21**	.20**	.46**	.47**	.61**	.60**
	Internalizing scale	.78**	.75**	.64**	.59**	.59**	.49**	.56**	.52**	.33**	.28**	.62**	.62**	.79**	.77**
CBCL[b]	Withdrawn	.55**	.54**	.37**	.36**	.39**	.31**	.55**	.51**	.27**	.25**	.29**	.23**	.53**	.51**
	Somatic complaints	.40**	.38**	.28**	.25**	.36**	.24**	.37**	.35**	.16	.10	.35**	.32**	.41**	.38**
	Anxious/ depressed	.64**	.64**	.45**	.44**	.55**	.42**	.53**	.48**	.33**	.26**	.36**	.28**	.62**	.58**
	Social problems	.33**	.31**	.17*	.15	.31**	.24**	.25**	.20*	.26**	.21*	.20*	.13	.33**	.28**
	Thought	.48**	.47**	.33**	.30**	.44**	.36**	.44**	.42**	.31**	.24**	.33**	.27**	.50**	.47**
	Attention	.44**	.43**	.23**	.21*	.48**	.41**	.45**	.41**	.22**	.17*	.27**	.23**	.45**	.43**
	Rule breaking	.49**	.45**	.32**	.27**	.48**	.42**	.39**	.37**	.20**	.17*	.33**	.31**	.48**	.45**
	Aggressive behavior	.35**	.29**	.22*	.16	.44**	.39**	.33**	.30**	.23**	.20*	.17*	.14	.37**	.34**
	Externalizing scale	.50**	.45**	.33**	.28**	.50**	.44**	.43**	.40**	.23**	.19*	.30**	.28**	.49**	.46**
	Internalizing scale	.65**	.61**	.45**	.43**	.53**	.42**	.57**	.53**	.32**	.27**	.42**	.35**	.64**	.60**
SCIM[c]	Total	.63**	.57**	.57**	.54**	.47**	.37**	.45**	.45**	.41**	.35**	.62**	.58**	.71**	.67**
	Consolidated	.54**	.50**	.40**	.38**	.31**	.17	.33**	.32**	.55**	.46**	.63**	.59**	.62**	.57**
	Disturbed	.42**	.39**	.43**	.41**	.40**	.37**	.37**	.36**	.20*	.16	.38**	.37**	.49**	.48**
	Lack of Self	.77**	.73**	.61**	.57**	.53**	.41**	.47**	.43**	.40**	.36**	.70**	.65**	.78**	.75**
Self-Harm		.57**	.56**	.43**	.39**	.44**	.38**	.35**	.33**	.25**	.25**	.48**	.50**	.57**	.57**

*p<.05, **p<.01

[a] correlations between DERS and YSR are for Samples 1, 2, and 3

[b] correlations between DERS and CBCL are for Samples 1 and 3

[c] correlations between DERS and SCIM are for Samples 2 and 3

Confirmatory Factor Analysis of the DERS-SF Confirmatory Factor Analyses (CFAs) were computed for the DERS-SF, using items that were identified in the process described above, and compared to the factor loadings of the original DERS. Analyses were conducted using Mplus 7.11 software (Muthén and Muthén 2013) using the maximum likelihood estimation (ML) procedure. ML is the most commonly used method for CFA, and provided a good fit with the data in this study since data were considered continuous, independent, and normally distributed. In addition, DERS items were rated on a 5-point scale, which exceeds the 2–3 category threshold in which WLSMV estimation become a preferred approach (Beaaducel and Herzberg 2006).

In our CFA models, each item was only allowed to load on one factor (i.e., cross-loadings were not estimated to improve model fit) to ensure that factor loadings were comparable across the DERS and the DERS-SF. The DERS-SF measurement model yielded good fit with the data (χ^2 (120)= 245.695, $p<0.01$; CFI=.96; TLI=.94; RMSEA=.06 (90 %: .05–.08); SRMR=.05) and item loadings ranged from .73 to .93, all well above acceptable levels. The CFA for the DERS

yielded marginal to poor fit with the data (χ^2 (579)= 1589.348, $p<0.01$; CFI=.85; TLI=.83; RMSEA=.08 (90 %: .08–.09); SRMR=.08). Factor loadings for items were very similar to those found for the DERS-SF, indicating consistency in the estimation of the latent factors for the DERS-SF and original DERS. Factor loadings are reported in Table 3.

We then computed within-measure correlations among subscales for the DERS and DERS-SF (see Table 4). Subscale correlations for the DERS-SF ranged from .04 to .85 and correlations among DERS scales ranged from .14 to .90, indicating similar performance for the two scales.

Psychometric Properties and Correlations Between the DERS-SF and DERS The psychometric properties of the DERS-SF and original DERS are reported in Table 5. Chronbach's alpha coefficients for each of the 3-item DERS-SF subscales exceeded .70 and ranged from .79 to .91. The DERS-SF values were comparable to the original DERS.

Correlations between DERS and DERS-SF subscales were calculated to evaluate the degree to which the two versions were similar for the participants. All correlations were above

Table 7 Comparisons of concurrent validity for the DERS and DERS-SF for college sample

		Strategies		Non-Accept		Impulse		Goals		Awareness		Clarity		Total	
		Full	SF	Full	SF	Full	SF	Full	SF	Full	SF	Full	SF	Full	SF
BDI-II		.70**	.66**	.51**	.48**	.56**	.49**	.52**	.49**	.28**	.28**	.50**	.48**	.68**	.66**
SCL-90-R	Obsessive Compulsive	.61**	.58**	.51**	.48**	.47**	.43**	.45**	.43**	.20**	.22**	.41**	.42**	.59**	.58**
	Interpersonal Sensitivity	.53**	.48**	.49**	.48**	.46**	.44**	.41**	.39**	.23**	.23**	.43**	.41**	.56**	.55**
	Depression	.60**	.57**	.53**	.50**	.51**	.47**	.44**	.42**	.20**	.21**	.42**	.42**	.60**	.59**
	Anxiety	.55**	.53**	.49**	.46**	.45**	.43**	.42**	.41**	.19**	.19**	.36**	.37**	.55**	.55**
	Hostility	.56**	.52**	.51**	.50**	.44**	.43**	.45**	.43**	.26**	.25**	.45**	.43**	.59**	.58**
	Phobic Anxiety	.57**	.54**	.51**	.49**	.46**	.45**	.47**	.46**	.20**	.19**	.39**	.39**	.58**	.58**
	Paranoid Ideation	.56**	.53**	.48**	.47**	.48**	.47**	.42**	.41**	.22**	.23**	.39**	.38**	.56**	.56**
	Psychoticism	.60**	.57**	.50**	.48**	.48**	.45**	.48**	.46**	.22**	.23**	.43**	.42**	.60**	.60**
	Somatization	.57**	.54**	.50**	.48**	.46**	.42**	.44**	.42**	.19**	.21**	.43**	.42**	.58**	.57**
	GSI	.64**	.60**	.56**	.54**	.52**	.48**	.49**	.48**	.22**	.23**	.45**	.45**	.64**	.64**
STAI-Y	State	.64**	.59**	.49**	.49**	.52**	.44**	.43**	.38**	.42**	.34**	.53**	.51**	.68**	.66**
	Trait	.76**	.70**	.59**	.56**	.60**	.49**	.49**	.43**	.46**	.35**	.59**	.58**	.79**	.75**
AAQ-II		−.78**	−.72**	−.61**	−.58**	−.68**	−.58**	−.51**	−.45**	−.44**	−.34**	−.58**	−.58**	−.82**	−.79**
SCIM	Total	.58**	.47**	.49**	.46**	.52**	.45**	.45**	.43**	.44**	.43**	.66**	.60**	.68**	.64**
	Consolidated	.35**	.28**	.23**	.21**	.27**	.24**	.28**	.24**	.40**	.38**	.46**	.36**	.44**	.39**
	Disturbed	.42**	.32**	.41**	.38**	.43**	.39**	.36**	.35**	.31**	.29**	.50**	.50**	.53**	.51**
	Lack of Self	.64**	.56**	.53**	.49**	.52**	.44**	.47**	.45**	.33**	.35**	.61**	.57**	.68**	.65**
Self-Harm		.27**	.23**	.24**	.23**	.20**	.18**	.18**	.20**	.09**	.12**	.17**	.15**	.25**	.25**

**$p<.01$, *$p<.05$

Correlations between DERS and BDI, Self-Harm, SCL are for both college samples; correlations between DERS and STAI, DERS and AAQ are only for Sample 1; correlation between DERS and SCIM are only for Sample 2

.90 and ranged from .91 to .98, indicating 83–96 % shared variance between the DERS-SF and original DERS scales, despite a drastic reduction in items.

Concurrent Validity of the DERS-SF Correlations were computed for the DERS and DERS-SF with several common outcome variables to allow for comparisons of the concurrent validity for the DERS and DERS-SF and provide information about how well the DERS-SF can be used to approximate the original version. Correlations for the DERS and DERS-SF with several outcome variables, including the CBCL, YSR, SCIM, and self-harm history are presented in Table 6. The correlations for DERS and DER-SF with indices of psychopathology were generally similar in patterns of statistical significance and in magnitude of correlation.

Study 2: Validation of the DERS-SF With a College Sample

Descriptive statistics for all self-report measures in Study 2 are summarized in Table 2.

Confirmatory Factor Analysis of the DERS-SF Just as was done with the adolescent sample, CFAs were computed for the DERS-SF and DERS with the college student sample. Consistent with the results for adolescents, the DERS-SF CFA yielded good fit with the data (χ^2 (120)=383.78, $p<0.01$; CFI=.97; TLI=.96; RMSEA=.05 (90 %: .05–.06); SRMR=.04). Factor loadings for all items were large, ranging from .63 to .90. The DERS CFA yielded marginal to poor fit with the data (χ^2 (579)=2825.61, $p<0.01$; CFI=.88; TLI=.87; RMSEA=.07 (90 %: .07–.07); SRMR=.07). However, as shown in Table 3, the factor loadings for the original DERS and the DERS-SF were similar, indicating good consistency in the estimation of the latent factors.

Intercorrelations among subscales within the DERS and DERS-SF were computed. As in Study 1, they were similar in statistical significance and magnitude (see Table 4). Correlations among DERS-SF scales ranged from .11 to .82. Correlations among the DERS scales ranged from .15 to .90. These findings indicate similar performance for the two scales.

Basic Psychometric Properties and Correlations Between the DERS-SF and DERS Table 5 reports the descriptive statistics and reliability of the DERS and DERS-SF. Consistent with the adolescent sample, the Chronbach's alpha coefficients for the DERS-SF total scale and six subscales all exceeded .70 and ranged from .78 to .91. Correlations between the DERS and DERS-SF again indicated strong correspondence in the two versions (Table 5). These correlations ranged from .90 to .97 and indicated that the DERS and the DERS-SF shared 81–94 % of their variance.

Concurrent Validity of the DERS-SF We calculated correlations between the original DERS, the short form, and their subscales. Concurrent validity of the DERS-SF was compared to the DERS, over 6 outcomes, as measured by the BDI-II, SCL-90-R, STAI-Y, AAQ-II, SCIM, and self-harm history (see Table 7). Similar to the adolescent sample, the patterns of correlations for the DERS and DERS-SF with these outcome variables were generally consistent in statistical significance and magnitude of correlation, across all scales. These findings suggest that the DERS-SF has comparable concurrent validity to the original DERS.

Discussion

This study was designed to evaluate whether a shortened version of the widely-used Difficulties in Emotion Regulation Scale can perform similarly to the full measure (Gratz and Roemer 2004). Consistent with our first hypothesis, results from the two confirmatory factor analyses presented here indicate the DERS-SF has sound psychometric properties that are comparable to or better than the original measure. Also consistent with this hypothesis, scores on the DERS-SF effectively capture the dimensions of emotion regulation deficits measured by the original DERS. Following from our second hypothesis, we found that correlations between scores on the DERS-SF and on other clinically relevant scales mirrored correlations observed when using the full DERS. As with previous research, DERS scores correlated moderately to highly with measures of both internalizing and externalizing psychopathology (see Tables 6 and 7). Finally, these results were consistent across both adolescent and adult samples, lending support to our third hypothesis that the DERS-SF may be useful for measuring emotion regulation deficits with participants across a wide age range. These findings supply further evidence that emotion regulation difficulties are an excellent transdiagnostic marker of psychopathology risk (Hofmann et al. 2012; Beauchaine and Thayer 2015).

Developing an abridged version of this widely used measure may facilitate research and enhance clinical practice aimed at targeting emotion regulation deficits. If higher reliability can be achieved from shorter tests, measurement error is reduced, statistical power enhanced, and inferences strengthened (Wilmer et al. 2012). Brief measures can maximize participant response rates and reduce response burden. Factors such as instrument length, cognitive effort required to complete a questionnaire, or survey layout and interface have been suggested to affect respondent strain (US Department of Health and Human Services, 2009). Further, lengthy questionnaires have been identified as a general obstacle in clinical practice (Mark et al. 2008) and instrument length has been used as an argument for limiting the overall number of

administrations of an instrument in longitudinal studies (Rolstad et al. 2011).

Strengths of the current studies include our relatively large pooled sample sizes, the inclusion of participants from diverse settings (i.e., community, clinical, inpatient, outpatient, and two different regions of the United States), and the wide range of ages represented across the 5 samples. There are also a number of limitations to the current study. We did not use the same outcome measures across the 5 samples, making it challenging to compare participants' scores systematically. The concurrent validity of the DERS has also been examined across a number of labs and with a broader range of outcome measures than those included in the current study. Furthermore, we lacked diagnostic interview data on study participants, which also limited our generalizability to clinical samples. Additional research will be needed to examine if the DERS-SF is a valid replacement for the original DERS when it is included in a longer battery. Although we were able to include participants from a range of developmental stages, certain groups are better represented than others. Our samples were predominantly Caucasian and most participants in our college sample were young adults. Thus, the utility of the DERS-SF needs further examination for use among those from more diverse racial/ethnic backgrounds, as well as persons in mid-to-late adulthood and adults outside of a college setting. There were also variations in the inclusion and exclusion criteria used to recruit our samples and females were overrepresented relative to males. Finally, we did not include other measures of emotion regulation difficulties that would have allowed us to examine convergent validity. Thus, although the DERS-SF performed similarly to the DERS, we do not yet know how it will compare to other gold-standard instruments of emotion regulation deficits.

In summation, the DERS-SF maintains the excellent psychometric properties and structure of the original DERS with half of the total items. This streamlined version of the instrument should be faster for participants and clients to complete and easier for researchers and clinicians to score. Given the frequent and broad used the DERS, we expect an abridged version to be useful in a range of settings, particularly those in which respondent time is limited and/or burden is high (e.g., epidemiological studies). Future studies should investigate the utility of the DERS-SF as a clinical outcome measure and examine if scores are sensitive to therapeutic change. Emotion regulation difficulties are important constructs for mental health practitioners and researchers interested in personality, psychopathology, and development. Efficient measurement of these constructs may be especially helpful for understanding the emergence and developmental course of several distinct forms of psychopathology.

References

Achenbach, T. M. (1987). *Manual for the youth self-report and profile*. Burlington: University of Vermont Department of Psychiatry.

Achenbach, T. M. (1991a). *Manual for the youth self-report and profile*. Burlington: University of Vermont Department of Psychiatry.

Achenbach, T. M. (1991b). *Manual for the child behavior checklist/4-18 and 1991 profile*. Burlington: University of Vermont Department of Psychiatry.

Achenbach, T. M., Edelbrock, C., & Howell, C. T. (1987). Empirically based assessment of the behavioral/emotional problems of 2- and 3-year-old children. *Journal of Abnormal Child Psychology, 15*, 629–650. doi:10.1007/BF00917246.

Barnes, L. L. B., Harp, D., & Jung, W. S. (2002). Reliability generalization of scores on the Spielberger state-trait anxiety inventory. *Educational and Psychological Measurement, 62*, 603–618. doi:10.1177/0013164402062004005.

Beauchaine, T. P., & Thayer, J. F. (2015). Heart rate variability as a transdiagnostic biomarker of psychopathology. *International Journal if Psychophysiology*. doi:10.1016/j.ijpsycho.2015.08.004.

Beauchaine, T. P., Gatzke-Kopp, L., & Mead, H. K. (2007). Polyvagal theory and developmental psychopathology: emotion dysregulation and conduct problems from preschool to adolescence. *Biological Psychology, 74*, 174–184. doi:10.1016/j.biopsycho.2005.08.008.

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling, 13*, 186–203.

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. (1996). Comparison of beck depression inventories -IA and -II in psychiatric outpatients. *Journal of Personality Assessment, 67*, 588–597. doi:10.1207/s15327752jpa6703_13.

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., & … Zettle, R. D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire–II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy, 42*, 676–688. 10.1016/j.beth.2011.03.007

Cappadocia, M., Desrocher, M., Pepler, D., & Schroeder, J. H. (2009). Contextualizing the neurobiology of conduct disorder in an emotion dysregulation framework. *Clinical Psychology Review, 29*, 506–518. doi:10.1016/j.cpr.2009.06.001.

Cole, P. M. (2014). Moving ahead in the study of the development of emotion regulation. *International Journal of Behavioral Development, 38*, 203–207. doi:10.1177/0165025414522170.

Cole, P. M., Michel, M. K., & Teti, L. O. (1994). The development of emotion regulation and dysregulation: A clinical perspective. In N. A. Fox (Ed.), The development of emotion regulation: Biological

and behavioral considerations. Monographs of the Society for Research in *Child Development*, 59 (Serial No. 240), 73–100.

Côté, G., Gosselin, P., & Dagenais, I. (2013). Évaluation multidimensionnelle de la régulation des émotions: propriétés psychométriques d'une version francophone du difficulties in emotion regulation scale psychometric properties of a french version of the difficulties in emotion regulation scale. *Journal de Thérapie Comportementale et Cognitive, 23*, 63–72. doi:10.1016/j.jtcc.2013.01.005.

Crowell, S. E., Beauchaine, T. P., McCauley, E., Smith, C., Stevens, A. L., & Sylvers, P. D. (2005). Psychological, physiological, and serotonergic correlates of parasuicidal behavior among adolescent girls. *Development and Psychopathology, 17*, 1105–1127. doi:10.1017/S0954579405050522.

Crowell, S. E., Baucom, B. R., Yaptangco, M., Bride, D., Hsiao, R., McCauley, E., & Beauchaine, T. P. (2014). Emotion dysregulation and dyadic conflict in depressed and typical adolescents: evaluating concordance across psychophysiological and observational measures. *Biological Psychology*. doi:10.1016/j.biopsycho.2014.02.009.

Derogatis, L. R., & Lazarus, L. (1994). SCL-90—R, brief symptom inventory, and matching clinical rating scales. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 217–248). Hillsdale: Lawrence Erlbaum Associates, Inc.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: tiny-yet-effective measures of the Big five factors of personality. *Psychological Assessment, 18*, 192–203. doi:10.1037/1040-3590.18.2.192.

Dvorak, R. D., Sargent, E. M., Kilwein, T. M., Stevenson, B. L., Kuvaas, N. J., & Williams, T. J. (2014). Alcohol use and alcohol-related consequences: associations with emotion regulation difficulties. *The American Journal of Drug and Alcohol Abuse, 40*, 125–130. doi:10.3109/00952990.2013.877920.

Edwards, P., Roberts, I., Clarke, M., DiGuiseppi, C., Pratap, S., Wentz, R., & Kwan, I. (2002). Increasing response rates to postal questionnaires: systematic review. *British Medical Journal, 324*, 1183–1185. doi:10.1136/bmj.324.7347.1183.

Eisenberg, N., & Spinrad, T. L. (2004). Emotion-related regulation: sharpening the definition. *Child Development, 75*, 334–339. doi:10.1111/j.1467-8624.2004.00674.x.

Folk, J. B., Zeman, J. L., Poon, J. A., & Dallaire, D. H. (2014). A longitudinal examination of emotion regulation: pathways to anxiety and depressive symptoms in urban minority youth. *Child and Adolescent Mental Health*. doi:10.1111/camh.12058.

Fossati, A., Feeney, J., Maffei, C., & Borroni, S. (2014). Thinking about feelings: affective state mentalization, attachment styles, and borderline personality disorder features among Italian nonclinical adolescents. *Psychoanalytic Psychology, 31*, 41–67. doi:10.1037/a0033960.

Gadow, K. D., Sprafkin, J., Carlson, G., Schneider, J., Nolan, E. E., Mattison, R. E., & Rundberg-Rivera, V. (2002). A DSM-IV-referenced adolescent self-report rating scale. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*, 671–679.

Gratz, K. L. (2001). Measurement of deliberate self-harm: preliminary data on the deliberate self-harm inventory. *Journal of Psychopathology and Behavioral Assessment, 23*, 253–263. doi:10.1023/A:1012779403943.

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, 26*, 41–54. doi:10.1023/B:JOBA.0000007455.08539.94.

Gratz, K. L., & Tull, M. T. (2010). The relationship between emotion dysregulation and deliberate self-harm among inpatients with substance use disorders. *Cognitive Therapy and Research, 34*, 544–553. doi:10.1007/s10608-009-9268-4.

Hayes, S. C., Strosahl, K., Wilson, K. G., Bissett, R. T., Pistorello, J., Toarmino, D., & … McCurry, S. M. (2004). Measuring experiential avoidance: a preliminary test of a working model. *The Psychological Record, 54*, 553–578.

Hofmann, S. G., Sawyer, A. T., Fang, A., & Asnaani, A. (2012). Emotion dysregulation model of mood and anxiety disorders. *Depression and Anxiety, 29*, 409–416. doi:10.1002/da.21888.

Jankowski, P. (2013). Identity status and emotion regulation in adolescence and early adulthood. *Polish Psychological Bulletin, 44*, 288–298. doi:10.2478/ppb-2013-0032.

Jankowski, P., & Rękosiewicz, M. (2013). Type of social participation and emotion regulation among upper secondary school students. *Polish Psychological Bulletin, 44*, 322–330. doi:10.2478/ppb-2013-0035.

Kaufman, E. A., Cundiff, J. M., & Crowell, S. E. (2015). The development, factor structure, and validation of the self-concept and identity measure (SCIM): A self-report assessment of clinical identity disturbance. *Journal of Psychopathology and Behavioral Assessment, 9*, 122–133. doi:10.1007/s10862-014-9441-2.

Kökönyei, G., Urbán, R., Reinhardt, M., Józan, A., & Demetrovics, Z. (2014). The difficulties in emotion regulation scale: factor structure in chronic pain patients. *Journal of Clinical Psychology, 70*, 589–600. doi:10.1002/jclp.22036.

Lavender, J. M., Wonderlich, S. A., Peterson, C. B., Crosby, R. D., Engel, S. G., Mitchell, J. E., & … Berg, K. C. (2014). Dimensions of emotion dysregulation in bulimia nervosa. *European Eating Disorders Review, 22*, 212–216. 10.1002/erv.2288

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York: Guilford Press.

Linehan, M. M., & Comtois, K. A. (1996). Lifetime parasuicide count: Unpublished Manuscript.

Mark, T., Johnson, G., Fortner, B., & Ryan, K. (2008). The benefits and challenges of using computer-assisted symptom assessments in oncology clinics: Results of a qualitative assessment. *Technology in Cancer Research & Treatment, 7*, 401–406. doi:10.1177/153303460800700508.

McLaughlin, K. A., Hatzenbuehler, M. L., Mennin, D. S., & Nolen-Hoeksema, S. (2011). Emotion dysregulation and adolescent psychopathology: a prospective study. *Behaviour Research and Therapy, 49*, 544–554. doi:10.1016/j.brat.2011.06.003.

Mirzaei, M., Gharraee, B., & Birashk, B. (2014). The role of positive and negative perfectionism, self-efficacy, worry and emotion regulation in predicting behavioral and decisional procrastination. *Iranian Journal of Psychiatry and Clinical Psychology, 19*, 230–240.

Mitchell, J. T., Robertson, C. D., Anastopolous, A. D., Nelson-Gray, R. O., & Kollins, S. H. (2012). Emotion dysregulation and emotional impulsivity among adults with attention-deficit/hyperactivity disorder: results of a preliminary study. *Journal of Psychopathology and Behavioral Assessment, 34*, 510–519. doi:10.1007/s10862-012-9297-2.

Muthén, L. K., & Muthén, B. O. (2013). *Mplus: statistical analysis with latent variables. User's guid (version 7.11)*. Los Angeles: Muthén & Muthén.

Neumann, A., van Lier, P. 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. doi:10.1177/1073191109349579.

Orgeta, V. (2009). Specificity of age differences in emotion regulation. *Aging & Mental Health, 13*, 818–826. doi:10.1080/13607860902989661.

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, 46*, 713–720. doi:10.1002/eat.22145.

Rolstad, S., Adler, J., & Rydén, A. (2011). Response burden and questionnaire length: is shorter better? a review and meta-analysis. *Value In Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research, 14*, 1101–1108. doi:10.1016/j.jval.2011.06.003.

Sarıtaş-Atalar, D., Gençöz, T., & Özen, A. (2014). Confirmatory factor analyses of the difficulties in emotion regulation scale (DERS) in a Turkish adolescent sample. *European Journal of Psychological Assessment*. doi:10.1027/1015-5759/a000199.

Sarıtaş-Atalar, D., Gençöz, T., & Özen, A. (2015). Confirmatory factor analyses of the difficulties in emotion regulation scale (DERS) in a Turkish adolescent sample. *European Journal of Psychological Assessment, 31*, 12–19. doi:10.1027/1015-5759/a000199.

Singh, P., & Singh, N. (2013). Difficulties in emotion regulation: a barrier to academic motivation and performance. *Journal of the Indian Academy of Applied Psychology, 39*, 289–297.

Smith, G. T., Combs, J. L., & Pearson, C. M. (2012). Brief instruments and short forms. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol 1: foundations, planning, measures, and psychometrics* (pp. 395–409). Washington, DC, US: American Psychological Association. doi:10.1037/13619-021.

Spielberger, C. D., Vagg, P. R., Barker, L., Donham, G., & Westberry, L. (1980). The factor structure of the state-trait anxiety inventory. *Stress and anxiety, 7*, 95–109.

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*, 385–392. doi:10.1016/j.janxdis.2012.01.005.

Thompson, R. A. (1994). Emotion regulation: A theme in search of definition. In N. A. Fox (Ed), The development of emotion regulation and dysregulation: Biological and behavioral aspects. *Monographs of the Society for Research in Child Development, 59* (2–3, Serial no. 240), 25–52.

US Department of Health and Human Services Food and Drug Administration; Center for Drug Evaluation and Research (CDER); Center for Biologics Evaluation and Research (CBER); Center for Devices and Radiological Health (CDRH), 2009. Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Available from: www.fda.gov/downloads/ Drugs/Guidance ComplianceRegulatoryInformation/Guidances/UCM193282.pdf. [Accessed October 2, 2014].

Vasilev, C. A., Crowell, S. E., Beauchaine, T. P., Mead, H. K., & Gatzke-Kopp, L. M. (2009). Correspondence between physiological and self-report measures of emotion dysregulation: A longitudinal investigation of youth with and without psychopathology. *Journal of Child Psychology and Psychiatry, 50*, 1357–1364. doi:10.1111/j.1469-7610.2009.02172.x.

Weinberg, A., & Klonsky, E. (2009). Measurement of emotion dysregulation in adolescents. *Psychological Assessment, 21*, 616–621. doi:10.1037/a0016669.

Weiss, N. H., Tull, M. T., Anestis, M. D., & Gratz, K. L. (2013). The relative and unique contributions of emotion dysregulation and impulsivity to posttraumatic stress disorder among substance dependent inpatients. *Drug and Alcohol Dependence, 128*, 45–51. doi:10.1016/j.drugalcdep.2012.07.017.

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Gerbasi, M., & Nakayama, K. (2012). Capturing specific abilities as a window into human individuality: the example of face recognition. *Cognitive Neuropsychology, 29*, 360–392. doi:10.1080/02643294.2012.753433.