



# Factor Structure of the Difficulties in Emotion Regulation Scale in Treatment Seeking Adults with Eating Disorders

Line Nordgren<sup>1</sup> · Elin Monell<sup>2</sup> · Andreas Birgegård<sup>2</sup> · Johan Bjureberg<sup>2,3</sup> · Hugo Hesser<sup>1</sup>

Published online: 4 December 2019

© The Author(s) 2019

## Abstract

The Difficulties in Emotion Regulation Scale (DERS) is extensively used as a measure of emotion (dys-)regulation ability in both clinical and nonclinical populations. This is the first study to examine the factor structure of both the original 36-item and short 16-item version of the DERS in adults with eating disorders and to test measurement invariance across diagnostic subgroups. The factor structure of the scale was examined using confirmatory factor analysis in a psychiatric sample of adults with eating disorders ( $N = 857$ ). Four primary factor structures were fitted to the data: (1) a unidimensional model, (2) a six-factor correlated-traits model, (3) a higher-order factor solution, and (4) a bifactor model. Measurement invariance was tested for diagnostic subgroups of anorexia nervosa and bulimia nervosa and associations between factors and eating pathology were examined in each diagnostic group. Results indicated that a modified bifactor solution fitted the data adequately for both the 36-item and 16-item version of the DERS. A general factor explained most of the variance (86%) and reliability was high for the general factor of DERS (total) but lower for the subscales. Measurement invariance of the bifactor model was supported across diagnostic subgroups and test of factor means revealed that bulimia nervosa had a higher factor mean than anorexia nervosa on the general factor. The general factor accounted for a significant proportion of variance in eating pathology. Our results support the use of the total scale of both the 36-item and 16-item version among adults with eating disorders.

**Keywords** Difficulties in emotion regulation scale · Emotion · Confirmatory factor analysis · Measurement invariance · Eating disorders

---

The data reported in this manuscript were obtained from Swedish Stepwise clinical database [STEPWISE/KÄTS]. A bibliography of journal articles using STEPWISE is available at [<http://www.atstorming.se/forskning-utbildning-2/forskning-inom-kats/kats-senaste-publicerade-forskning-om-atstormingar/>]. The main variable, Difficulties in Emotion Regulation Scale (DERS), examined in the present article has been used in one previous study (cited in the current manuscript). To the best of our knowledge the DERS has never been analyzed with the same statistical methods or with the same study aim as in the present study.

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10862-019-09765-8>) contains supplementary material, which is available to authorized users.

---

✉ Hugo Hesser  
hugo.hesser@liu.se

Line Nordgren  
line.nordgren@liu.se

Elin Monell  
elin.monell@ki.se

Andreas Birgegård  
Andreas.Birgegard@ki.se

Johan Bjureberg  
johan.bjureberg@ki.se

- <sup>1</sup> Department of Behavioural Sciences and Learning, Linköping University, SE-581 83 Linköping, Sweden
- <sup>2</sup> Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, Norra Stationsgatan 69, SE-113 64 Stockholm, Sweden
- <sup>3</sup> Stockholm Health Care Services, Stockholm County Council, Stockholm, Sweden

Eating disorders can be described as a prolonged disturbance of eating or related behavior such as selective or restrictive eating, compulsory eating behavior, binge eating episodes and/or compensatory behavior such as fasting, frequent vomiting, use of laxatives or diuretics as well as excessive and compulsory exercise (American Psychiatric Association 2013). Eating disorders are divided into sub-diagnoses anorexia nervosa, bulimia nervosa, binge eating disorder and other specified feeding or eating disorders, the latter including atypical anorexia nervosa, bulimia nervosa and binge eating disorder of low frequency and/or limited duration, rumination disorder, and purging disorder (American Psychiatric Association 2013). Overall eating disorder point prevalence has been reported between 0.5% and 5.3% for females and from 0.62% to 0.64% in males (Lindvall Dahlgren and Wisting 2016). In eating disorders, emotion regulation seems to play an important role. Previous research, for example, indicates that the risk of suicide attempts is highly elevated in individuals with eating disorders (Pisetsky et al. 2013) and a link between suicide attempts, non-suicidal self-injury and emotion dysregulation in individuals with eating disorders has been suggested (Gómez-Expósito et al. 2016; Vieira et al. 2016).

Emotion regulation has emerged as a central component of both theories of psychopathology and psychological interventions (Gross 2015). Emotion regulation can be conceptualized as the awareness, understanding and acceptance of one's emotions, the ability to inhibit inadequate behaviors when emotionally aroused, and the ability to use adaptive regulation strategies in order to reach one's goals (Gratz and Roemer 2004). Individuals with eating disorders have been found to display higher levels of emotion dysregulation compared to healthy controls, lower levels of emotional awareness, clarity, and recognition as well as problems regarding emotional inhibition and access to healthy emotion regulation strategies (e.g., Lavender et al. 2015; Monell et al. 2018). Regarding different eating disorder diagnostic subgroups, Lavender et al. (2015) concludes in a review that global emotion regulation difficulties seems to be a transdiagnostic trait across the eating disorder spectrum, but that there might be some distinct patterns that potentially distinguishes the different eating disorders. Although numerous studies have reported relationships between eating disorders and emotion dysregulation (e.g., Brockmeyer et al. 2014; Lavender et al. 2015), a consistent conceptualization of emotion regulation seems to be lacking within the field of eating disorders. This is evidenced by the use of a wide range of different measures such as the subscale interoceptive awareness from the Eating Disorder Inventory (Garner et al. 1983), the Emotional Awareness Questionnaire (Rieffe et al. 2007), the Toronto Alexithymia Scale (Bagby et al. 1994), the Emotion Regulation Questionnaire (Gross and John 2003) and the Difficulties in Emotion Regulation Scale (DERS; Gratz and Roemer 2004).

Given the growing interest in emotion regulation in eating disorders, there is a need for valid and reliable measures. One widely used measure in both research and clinical practice is the aforementioned DERS (Gratz and Roemer 2004). Unlike most other measures that mainly focus on specific aspects of emotion regulation, the DERS was designed to comprehensively measure multiple dimensions of emotion regulation ability. The original version of the DERS includes 36 items formulated as assertions scored 1–5 yielding a total score as well as scores on six subscales (henceforth referred to as DERS-36). In the original study by Gratz and Roemer (2004), exploratory factor analysis (EFA) revealed a correlated trait lower-order six-factor solution. The six factors or subscales were presented as distinct but related dimensions with adequate internal consistency. All items in the final exploratory factor solution had factor loadings of .40 or higher on the corresponding subscale and none of the items in the final version had significant loadings (above .40) on more than one factor (Gratz and Roemer 2004).

The six subscales were named *Nonacceptance*, *Goals*, *Impulse*, *Awareness*, *Strategies*, and *Clarity*. The Nonacceptance subscale (nonacceptance of emotional responses) is composed of items reflecting a tendency to experience negative secondary emotional responses or a nonaccepting reaction to distress, mainly shame, guilt, or self-blame regarding one's own (negative) emotions. Items from the Goals subscale (difficulties engaging in goal directed behavior) reflect difficulties regarding concentration or accomplishing tasks when upset. The Impulse subscale (impulse control difficulties) reflect difficulties remaining in control of one's behavior when feeling distress and also includes items reflecting the perception of emotions as overwhelming. The fourth subscale Awareness (lack of emotional awareness) comprises reverse scored items reflecting the tendency to pay attention to and acknowledge emotions. The subscale Strategies (limited access to emotion regulation strategies) consists of items reflecting a belief that there is nothing that can help regulate negative emotions, suggesting a kind of hopelessness when confronted with one's feelings. The final subscale Clarity (lack of emotional clarity) addresses the ability to understand emotions; a high score indicates a high degree of confusion regarding emotions.

Previous research has demonstrated support for the validity of the DERS. The DERS and its subscales has shown correlations with other scales measuring related aspects of emotion regulation, such as alexithymia, levels of positive and negative affect, experiential avoidance and suppression, in other populations including individuals with alcohol dependence, chronic pain, and aggression as well in psychiatric inpatient adolescents (e.g., Ghorbani et al. 2017; Kökönyei et al. 2014; Velotti et al. 2016; Venta et al. 2012). Levels of emotion regulation difficulties as measured by the DERS have also been found to be associated with different kinds of psychopathology such as

personality disorders, posttraumatic stress disorders, depression and anxiety (e.g., Gratz et al. 2006; Sloan et al. 2017; Tull and Roemer 2007; Villalta et al. 2018; Visted et al. 2018). Regarding eating disorders, previous research has found associations between the DERS (total scale and all subscales) and levels of eating disorder pathology (e.g., Brockmeyer et al. 2014; Burns et al. 2012; Cooper et al. 2014; Haynos et al. 2015; Monell et al. 2015; Racine and Wildes 2015), and there is also evidence of relations to specific eating disorder behaviors (e.g., Burns et al. 2012; Monell et al. 2018). Furthermore, studies incorporating other emotion regulation measures in addition to DERS show that the relationships between the DERS and its subscales and levels of eating disorder pathology are fairly consistent with associations that were observed when using other measures that tap into similar constructs (e.g., Svaldi et al. 2012). Levels of emotion regulation difficulties in individuals with eating disorders, as measured by the DERS, have also been found to decrease following treatment (e.g., Mallorquí-Bagué et al. 2018; Sloan et al. 2017).

Since the original factor analytic study, a number of studies have examined the psychometric properties of the DERS-36. The majority of these studies have, however, been conducted in non-clinical samples. Several of the previous studies report acceptable or good fit for the original correlated traits six-factor solution in non-clinical adults (e.g., Bostan and Zaharia 2016; Ritschel et al. 2015) and adolescents (e.g., Neumann et al. 2010; Saritaş-Atalar et al. 2015). However, a few studies investigating the factor structure of the DERS-36 in non-clinical samples report trouble with the original six-factor solution, primarily relating to problems with the Awareness subscale (e.g., Bardeen et al. 2012; Lee et al. 2016). For example, Bardeen et al. (2012) found support for a revised five-factor higher order solution and Lee et al. (2016) suggested a five-factor lower order model – both excluded the six items from the Awareness subscale.

While the factor structure of the DERS-36 has been examined in non-clinical samples, only a handful of studies have examined the factor structure in clinical populations (Fowler et al. 2014; Hallion et al. 2018; Osborne et al. 2017; Perez et al. 2012; Wolz et al. 2015). Perez et al. (2012) investigated the factor structure of DERS-36 in a sample of 218 adolescents with nonsuicidal self-injury. Confirmatory factor analysis (CFA) suggested acceptable fit for the six-factor original correlated traits solution. Fowler et al. (2014) examined a sample of 592 adult in-patients with severe mental illness and reported acceptable and equivalent fit with the original correlated traits six-factor solution. In the recent study by Osborne et al. (2017), the factor structure of the DERS-36 was examined in an adult population of 344 outpatients receiving Dialectical Behavior Therapy (DBT). Seven different models were tested, including two unidimensional models, three correlated traits models, one higher-order model and one bifactor model. Results showed good fit for the bifactor

model including the original six factors, with a modification that prohibited items from the Awareness subscale from loading on the general factor but allowing the factor Awareness to be correlated with the Clarity subscale (Osborne et al. 2017). Support for a bifactor solution was replicated by Hallion et al. (2018) using CFA in a study of 427 adults with emotional disorders. However, unlike the study by Osborne et al. (2017), Hallion et al. (2018) found support for a five-factor bifactor solution that excluded the subscale Awareness altogether.

The only study to our knowledge to examine the factor structure of the DERS-36 in patients with eating disorders, is the study by Wolz et al. (2015). The study involved 74 healthy controls and 134 adult patients with a DSM-IV eating disorder (including anorexia nervosa, bulimia nervosa, binge eating disorder, and other specified feeding or eating disorders). The sample was analyzed with both EFA and CFA. EFA results suggested a one-dimension solution as well as a six-factor solution with close resemblance to the original factor structure. CFA results showed acceptable fit for a model comparable to the original six-factor correlated traits solution (Wolz et al. 2015).

In addition to the original 36-item version of the DERS, Bjureberg et al. (2016) developed and evaluated a brief version of the DERS consisting of 16 items that generates a single global score for emotion regulation difficulties (henceforth referred to as the DERS-16). The DERS-16 showed good convergent and discriminant validity, excellent internal consistency and good test-retest reliability and was concluded to offer a valid and brief method for assessing emotion regulation difficulties (Bjureberg et al. 2016). The factor structure of the DERS-16 has been investigated in five studies. Miguel et al. (2017), Shahabi et al. (2018), Westerlund and Santtila (2018), and Yiugit and Guzey Yiugit (2017) all found support for the five-factor solution proposed by Bjureberg et al. (2016) in non-clinical populations. Hallion et al. (2018) examined the factor structure of the DERS-16 in a clinical sample consisting of 427 adults with emotional disorders. Results indicated good fit for a bifactor model.

In summary, the factor structure of DERS-36 and DERS-16 has been investigated in several studies. However, most factor analytic studies are based on non-clinical samples and there is limited knowledge of the factor structure of both versions in clinical samples. To our knowledge, only one study has examined the DERS-36 in an eating disorder population (Wolz et al. 2015). That study, however, was marked with some limitations, most significantly the relatively small sample size for the analyses that were conducted in the study (i.e., CFA). The small sample size also precluded a careful examination of the factor structure in different subpopulations of eating disorders. Further, only one model was investigated using CFA meaning that there were no direct comparisons between different factor solutions. Also, given that both EFA and CFA

was conducted using the same sample in the study, there is a clear need to confirm the factor structure using CFA in an independent larger sample of individuals with eating disorders.

The present study is the first to investigate the factor structure of both the 36-item and 16-item DERS in a large eating disorder sample. Based on previous examinations of the factor structure of the 36-item DERS in clinical and non-clinical samples, a series of CFA models were fitted and compared in the present study. Given previously identified problems with the Awareness subscale of the DERS-36, a priori modifications were made to each model that was tested. Specifically, four primary structures were fitted to the data and compared in terms of model fit: (1) a unidimensional model, (2) a six-factor correlated-traits model, (3) a higher-order factor model, and (4) a bifactor model. We were primarily interested in the bifactor model that has to date not been examined among individuals with eating disorder but has previously been successfully tested on the DERS-36 in two psychiatric samples. The bifactor model also has a clear advantage by examining the specific variance accounted for by the subscales of DERS-36 over and above the variance accounted for by the general factor. Thus, information gained from the bifactor model can, for example, be used to determine the adequacy of the total score of the original DERS-36 and the DERS-16, as well as the additional value of scoring subscales as the correlations among all the items are accounted for by both a general factor and set of subscales in the model. Finally, the present study is, to our knowledge, the first to test for measurement and structural invariance across subgroups of different eating disorders, that is, anorexia nervosa and bulimia nervosa.

## Methods

### Participants

The sample was drawn from the Swedish Stepwise clinical database (Birgegård et al. 2010) with nationwide data on treatment seeking individuals of all ages and both genders with eating disorder; however, in the present study only adult patients ( $\geq 18$  years) were included. Data was extracted in October 2017. Stepwise is a longitudinal internet-based quality assurance register developed in 2005. Inclusion criteria for entry in the Stepwise database are a DSM-IV eating disorder (American Psychiatric Association 2000), medical or self-referral to a participating treatment unit, and intention to treat at the unit. All measures are collected during an initial assessment within the first three visits at the clinical unit. Assessment was carried out by eating disorder professionals with special training (a 2-day mandatory course). The Stepwise database includes structured diagnostic interviews

for eating disorders and other DSM-IV Axis I disorders and contains a variety of both mandatory and optional assessment instruments (i.e. units can decide which of the optional instruments to include in their test battery). For a full description of Stepwise, see Birgegård et al. (2010).

The DERS-36 was added as an optional assessment for units on 7th April 2014. Registrations prior to this date were thus excluded, leaving  $N = 2625$  patients from 30 units. The following exclusions were then made: no DERS ratings ( $n = 1674$ ), ratings from patients at units with  $< 20$  DERS-36 administrations ( $n = 94$ ), leaving a final sample size of 857 patients from 14 units. A previous study using the DERS-36 from the Stepwise register, largely sharing the present sample ( $N = 999$ , only women), could not find any differences between patients with or without DERS-36 registration regarding variables such as age, body mass index, eating disorder characteristics, and psychiatric comorbidity (see Monell et al. 2018).

The present sample consisted of 820 women and 37 men with a mean age of 26. As is typical in the eating disorder population, mean age of eating disorder onset was in adolescence. The majority of the participants had graduated from upper secondary school and had an employment of at least half-time. Bulimia nervosa was the most common diagnosis in the sample ( $n = 319$ ), followed by purging disorder ( $n = 189$ ), anorexia nervosa ( $n = 141$ ; whereof  $n = 90$  for restricting subtype,  $n = 32$  for binge/purging subtype, and  $n = 19$  for anorexia nervosa except no amenorrhea), atypical anorexia nervosa ( $n = 131$ ), binge eating disorder ( $n = 41$ ), unspecified feeding or eating disorder ( $n = 23$ ) and rumination disorder ( $n = 13$ ). Demographic information for the final sample is presented in Table 1. Eating disorder diagnoses were recategorized to correspond to the diagnostic criteria in DSM-5 (American Psychiatric Association 2013); eating disorders not otherwise specified type 1 (anorexia nervosa except no amenorrhea) were recategorized into anorexia nervosa and eating disorders not otherwise specified type 3 (bulimia nervosa except binge/purge frequency lower than the DSM-IV threshold) into bulimia nervosa.

### Measures

**DERS** DERS is a self-assessment scale measuring emotion dysregulation. The original DERS includes 36 items scored 1–5 where 1 is almost never (0–10%), 2 is sometimes (11–35%), 3 is about half the time (36–65%), 4 is most of the time (66–90%), and 5 is almost always (91–100%). Of the 36 items, 11 are reverse scored. The DERS-36 yields a total score as well as six subscales where higher scores indicate more difficulties. The DERS-36 has demonstrated adequate construct and predictive validity as well as good test-retest reliability (Gratz and Roemer 2004). The DERS-16 is a short version developed by Bjureberg et al. (2016). In this study,



**Table 1** Descriptive and clinical characteristics of the sample

Characteristics	Total (N = 857)			Anorexia Nervosa (n = 272)			Bulimia Nervosa (n = 319)		
	<i>M</i>	<i>SD</i>	% ( <i>n</i> )	<i>M</i>	<i>SD</i>	% ( <i>n</i> )	<i>M</i>	<i>SD</i>	% ( <i>n</i> )
Age	26.37	8.23		24.24	7.38		27.09	8.19	
Age of ED onset	16.45	5.60		16.83	5.56		15.65	4.05	
BMI	22.94	6.16		18.75	3.83		25.32	5.57	
EDE-Q	3.86	1.19		3.66	1.36		4.15	0.96	
Gender			96 (820)			94 (255)			97 (308)
Education			80 (688)			22 (60)			31 (100)
Employment			79 (673)			76 (207)			83 (264)
Living alone			39 (337)			34 (91)			42 (134)

The subgroup Anorexia Nervosa in this table includes a combination of anorexia nervosa ( $n = 141$ ) and atypical anorexia nervosa ( $n = 131$ ); BMI, Body Mass Index; EDE-Q, Eating Disorder Examination Questionnaire version 4.0, total score; Gender, female; Education, upper secondary graduate; Employment, employment above 50%; ED, eating disorder

the DERS-36 was administrated in Swedish to all participants and the items composing the DERS-16 were extracted from participants' scores on the original 36-item version. The Swedish DERS-36 and DERS-16 have been used in several previous studies with retained reliability and validity (e.g. Bjureberg et al. 2016; Garke et al. 2019; Monell et al. 2018; Monell et al. 2015). In the selected sample of participants, there were no missing data on the DERS.

**Structured Eating Disorder Interview (SEDI)** SEDI is a semi-structured clinical interview assessing eating disorder symptoms and establishes a preliminary DSM-IV eating disorder diagnosis which is then confirmed by an expert clinician. The interview includes 20–25 questions and is based on SCID-I (First et al. 2002). Preliminary validation against Eating Disorder Examination Interview (Cooper and Fairburn 1987) has shown concordance of 81% for specific eating disorder diagnosis and subtype (de Man Lapidoth and Birgegård 2010).

**Eating Disorder Examination Questionnaire (EDE-Q Version 4.0)** EDE-Q is a self-report measure assessing eating disorder symptomatology with 36 items scored 0 to 6, where 22 of the 36 items yields four subscales where higher scores indicate more pathology: Restraint, Eating concern, Shape concern, and Weight concern, as well as a mean Global score (Fairburn and Bèglin 1994). The EDE-Q is frequently used in eating disorder treatment and research and it has established reliability (Luce and Crowther 1999) and validity (Mond et al. 2004).

## Data Analytic Approach and Statistical Analysis

**Confirmatory Factor Analysis (CFA)** To examine the underlying factor-structure and to identify the best-fitted factor model, a series of CFA models were fitted based on previous

examinations of the factor structure of the 36-item DERS in clinical and non-clinical samples. Models were fitted with the non-normality robust maximum likelihood estimator (MLR) using Mplus vs. 7.4 (Muthén 2015).<sup>1</sup> The primary structures that were examined were (a) a unidimensional model (Model 1), (b) the original posited six-factor correlated-traits model (Model 2), (c) a second-order model with uncorrelated six-factors and a higher-order factor representing general emotion regulation ability (Model 3), and (d) a bifactor model in which items loaded directly on both the general factor and their corresponding subscale factor (Model 4). Given previously identified problems with the Awareness subscale of the DERS-36, one a priori modification was made to each one of these primary CFA structures (Model 1–4). The six Awareness items were either excluded from the analysis (model 1.2 and 2.2) or the Awareness factor was allowed to be correlated with the Clarity factor in models where covariance between factors were constrained to zero as default (model 3.2 and model 4.2). In total, eight models were fitted and compared to identify the best-fitted model. Once the best-fitted model was identified, the model was subsequently used to test for measurement and structural invariance across subgroups of participants with different eating diagnoses, for scale reliability analyses, and for test of the factor structure of the DERS-16.

To evaluate and compare model fit, the Root Mean Square Error of Approximation (RMSEA) were used with values

<sup>1</sup> We also reran the primary set of models in which we controlled for dependence in the data due to the nesting of individuals within clinics (using the Cluster option in Mplus) and with a different estimator (robust weighted least squares means and variance adjusted estimator [WLSMV]) that treated the indicators as ordered categorical. Although this improved model fit overall, these analyses did not alter the overall conclusions in terms of the best-fitted model among the set of estimated models. For this reason and given that the WLSMV estimator cannot produce estimates of error variance that is required for the computation of scales reliability statistics in the form of Omega and OmegaH, we present the results from analyses with MLR estimator that treated indicators as non-normal continuous.

<.08 and <.06 as benchmarks for adequate and good fit (Hu and Bentler 1999; Wang and Wang 2012). In addition, we provided the 90% confidence interval (CI) around the RMSEA value and only considered a model to have a good fit if the upper limit was below .08 (Wang and Wang 2012). Simulation studies have shown that RMSEA performs better than other fit indices in certain applications (see Wang and Wang 2012), and we thus relied on it as the primary measure of fit. We also employed the Comparative Fit Index (CFI) (Bentler 1990), with values >.90 and >.95 as benchmarks for adequate and good model fit (Bentler 1990; Hu and Bentler 1999), and the Standardized Root Mean Square Residual (SRMR) with values <.08 as a benchmark for a good fitted model (Hu and Bentler 1999). When comparing nested models, a scaled chi-square difference test was used for model comparisons (Muthén 2015; Satorra and Bentler 2001). A more constrained model (null model) was deemed to fit worse than an alternative less restrictive model if the increase in chi-square statistics was statistically significant at the level of  $p < .05$ , with degrees of freedom equal to the difference in free parameters between models. In addition, following recent recommendations for testing measurement invariance (Cheung and Rensvold 2002; Wang and Wang 2012), we also examined a drop in CFI between models with a delta CFI less than or equal to .01 as indicative of no significant difference between models.

#### Scale Reliability and Explained Common Variance (ECV)

Reliability measures for total scales and subscales were based on estimated loadings and measurement errors. Specifically, for both total scales and subscales, we calculated *Omega* (Raykov 1997) that uses all sources of reliable variances in the calculation of estimates of reliability (and is comparable to a weighted alpha coefficient; Reise 2012; Rodriguez et al. 2016a, b), and *Hierarchical Omega* (OmegaH; Rodriguez et al. 2016b; Zinbarg et al. 2005) that uses *only* the reliable variance that is specifically accounted for by the particular scale under examination (disregarding other reliable variance that was not accounted for by the scale). Similar to coefficient alpha, values range from 0 (no reliability) to 1 (perfect reliability). These “model-based” indices of reliability are generally considered to yield more accurate estimates of reliability than simple alpha coefficients mainly because they do not assume that each item is measured with the same degree of precision and measurement error (as assumed with coefficient alpha; Graham 2006; Raykov 1997). In addition, OmegaH has been suggested to be particularly useful for models that provide estimates of variance accounted for by *both* a general factor and subscale factor (e.g., bifactor or higher-order factor models) because these indices separate reliable variance explained by the general and subfactors and hence do not confound variance (Rodriguez et al. 2016b). Specifically, OmegaH for the total scale provides an estimate of the reliable variance

explained by all items loading on the general factor whilst partialling out reliable variance accounted for by the loadings on the subscales, whereas OmegaH for the subscale provides an estimate of the reliable variance explained by the items for the subscale whilst partialling out reliable variance accounted for by these items’ loadings on the general factor.

Based on recommendations for evaluating bifactor models, we also calculated ECV as a measure of how much of the common variance was explained by the general factor (Rodriguez et al. 2016b). High ECV indicates that the majority of the common variance is accounted for by the general factor, whereas low ECV indicates that most of the explained variance is accounted for by the subscales. ECV can hence be thought of as an indicator of the degree to which the construct is unidimensional (Rodriguez et al. 2016b).

**Measurement Invariance** Once a global structure had been identified in the full sample, we aimed to test for measurement invariance across groups with different eating diagnoses following recommendations provided by Wang and Wang (2012). Specifically, we tested measurement invariance of the best-fitted model across individuals with bulimia nervosa ( $n = 319$ ) and individuals with anorexia nervosa ( $n = 272$ ). Data from participants with atypical anorexia nervosa diagnosis and anorexia nervosa diagnosis were combined to form the anorexia nervosa group. This was done for three reasons. First, these diagnoses only differ in terms of the weight criteria (for atypical anorexia nervosa the individual has lost a significant amount of weight although still not considered significantly underweight) and thus we considered the groups to be similar in terms of diagnostic criteria. Second, this increased sample size to an acceptable level for these complex SEM models. Third, as reported in previous studies (e.g., Moskowitz and Weiselberg 2017; Sawyer et al. 2016), individuals with anorexia nervosa and atypical anorexia nervosa show more similarities than differences in levels of malnutrition and dietary restrictions, type of medical complications and psychological morbidities.

To establish factorial structure invariance, the best-fitted model identified in the full sample was first fitted in each subsample to examine whether the overall factor structure was similar across groups. The model fit indices used to evaluate model fit in the full sample were used to determine whether the model fit adequately in each subsample. The models were then combined into a multigroup CFA to test for both configural invariance and strong measurement invariance (i.e., invariant loadings and intercepts; also known as Scalar invariance; Meredith 1993). The configural model was also fitted as a baseline model in which subsequent more restricted models could be compared. The baseline model (i.e., configural model) included the same factor structure with the same pattern of fixed and free loadings in each group but without posing any equality restrictions on any measurement

(intercepts, loadings, and residuals) and structural parameters (factor variance) across groups. To formally test for strong measurement invariance (Meredith 1993), we then fitted a more restricted model in which both factor loadings and item intercepts were constrained to be invariant across groups and tested whether these constraints significantly degraded model fit. If model fit was not significantly worse in the constrained model as compared to the baseline model, strong measurement invariance was presumed to be established. Given that the measurement invariant model was nested within the configural model (i.e., baseline model), the scaled chi-square difference test and delta CFI were used to determine whether constraints significantly worsened model fit.

**Structural Invariance** Once measurement invariance was established, we proceeded to test the structural parameters of the best-fitted CFA model using multiple-group CFA analyses. Specifically, we first tested the factor variance and then proceeded to test the factor means (using models with equality constraints on loadings and intercepts in accordance with the strong measurement invariance assumption). Factor variance was set to be invariant in the model and this model was compared with the same model but where factor variances were free parameters in all groups. The scaled chi-square difference test and the decrement in CFI were used to examine whether equality constraints on factor variance significantly worsened model fit (Wang and Wang 2012).

Following the test of factor variance, we proceeded and tested whether the factor means for the factor of the subscales and the total scale differed as a function of eating disorder group (using the anorexia nervosa group as the reference category). In order not to inflate alpha for multiple comparisons, we used the scaled chi-square difference test (similar to ANOVA) to examine an overall difference between a model that constrained all factor means to be equal across groups with a model in which factor means differed as a function of eating disorder group (using one group as a reference category for identification purposes). If the ratio test was found to be statistically significant, we proceeded to test each factor mean difference with normal theory tests (i.e., estimate/standard error).

**Criterion Validity** In the best-fitted CFA model, we examined the contribution of each factor dimension in accounting for variability in eating pathology as measured with EDE-Q. Our aim was to explore whether each latent factor could account for unique variance in the external criterion by incorporating EDE-Q as a dependent variable in the structural part of the CFA model (similar to regression analysis). By using the latent factors, as compared to the observed scores of the scales, we could eliminate regression bias due to measurement error and explicitly model the true score variance of each factor and their diverging associations with the external

criterion. Models were run separately for individuals with bulimia nervosa ( $n = 319$ ) and individuals with anorexia nervosa ( $n = 272$ ).

## Results

### Factor Structure of the DERS-36

Descriptive statistics for each item of the DERS-36 are presented in supplementary Table S1 and descriptives for the total scale of DERS and subscales are presented in Table 2. Measures of skewness and kurtosis for both items and scales were generally low. In total, 8 models were estimated using CFA. Table 3 provides model fit indices for all estimated models in the full sample ( $N = 857$ ).

**Unidimensional Models** The unidimensional model (model 1.1) using all 36 items of the DERS-36 provided a poor fit to the data (see Table 3). The same was true for the second unidimensional model (model 1.2) that excluded the six Awareness items.

**Traits Models** The six-factor correlated traits model based on all 36 items of the DERS showed adequate fit according to RMSEA but only approached adequate CFI and SRMR (model 2.1). Factor loadings were all significant in the model. Correlations between latent factors were in the range of .20 to .75. Noteworthy is that the Awareness factor had relatively low correlations with all other factors (all  $<.30$ ) with the only exception for the correlation with the Clarity factor (.75). The five-factor correlated traits model that excluded the Awareness items evidenced acceptable model fit (model 2.2). All items had significant factor loadings and correlations between latent factors ranged from .45 to .78.

**Higher-Order Models** The first higher-order factor model that allowed the six factors to load on a higher-order factor evidenced acceptable fit according to RMSEA but only approached an acceptable CFI and SRMR (model 3.1). The modified version of the higher-order model (model 3.2), which allowed the Awareness and Clarity factor to be correlated, did improve model fit significantly,  $\Delta\chi^2(1) = 209.362$ ,  $p < .001$ . However, examining model fit indices indicated no substantial improvement in fit (see Table 3). Factor loadings were all significant in the model and the model estimated correlation between the Awareness and Clarity factor was .75.

**Bifactor Models** The first unmodified bifactor model that allowed items to load directly on both the general factor and each subscale factor (model 4.1) evidenced an acceptable fit according to RMSEA but only approached acceptable CFI

**Table 2** Descriptive statistics for total and subscales of Difficulties in Emotion Regulation Scale (36 items) as a function of diagnostic groups

Subscale	Total ( <i>N</i> = 857)						Anorexia Nervosa ( <i>n</i> = 272)						Bulimia Nervosa ( <i>n</i> = 319)					
	Sum	<i>SD</i>	<i>M</i>	<i>SD</i>	Skew	Kurt	Sum	<i>SD</i>	<i>M</i>	<i>SD</i>	Skew	Kurt	Sum	<i>SD</i>	<i>M</i>	<i>SD</i>	Skew	Kurt
Total	100.12	26.43	2.80	0.73	0.10	0.17	99.31	27.38	2.78	0.76	0.16	−0.32	102.32	26.62	2.86	0.73	−0.00	−0.53
Non-accept	16.15	6.22	2.69	1.04	0.37	−0.72	16.15	6.10	2.69	1.02	0.30	−0.69	16.17	6.34	2.69	1.06	0.39	−0.84
Goals	16.30	5.30	3.26	1.06	−0.20	−0.93	15.93	5.41	3.19	1.08	−0.08	−1.13	16.87	5.16	3.37	1.03	−0.39	−0.75
Impulse	14.39	6.17	2.40	1.03	0.61	−0.48	13.85	6.21	2.31	1.03	0.74	−0.30	15.11	6.15	2.52	1.02	0.46	−0.65
Awareness	18.23	5.15	3.04	0.86	−0.15	−0.55	18.24	5.15	3.04	0.86	0.02	−0.52	18.40	5.14	3.07	0.86	−0.26	−0.48
Strategies	21.10	7.75	2.64	0.97	0.29	−0.76	20.72	8.11	2.59	1.01	0.40	−0.70	21.76	7.64	2.72	0.95	0.18	−0.82
Clarity	13.95	4.59	2.79	0.92	0.15	−0.58	14.40	4.67	2.88	0.93	0.03	−0.65	14.01	4.61	2.80	0.92	0.16	−0.51

Skew, Skewness; Kurt, Kurtosis. The scores from DERS total scale and subscales are presented as sums as well as means, with the purpose to enable standardized comparisons between subscales due to the fact that the DERS subscales consist of different number of items (5–8 items per subscale)

and SRMR. However, the modified version of the bifactor model (model 4.2), which also estimated the correlation between the Awareness and Clarity factors, improved model fit,  $\Delta\chi^2(1) = 180.284, p < .001$ .<sup>2</sup> The model evidenced a good fit according to RMSEA and SRMR, and an acceptable fit according to CFI.

**Model Comparisons** Examination of fit indices provided the strongest support for the modified bifactor model (model 4.2). Specifically, whereas the modified correlated traits model was acceptable, the final bifactor model was the only model that evidenced a good fit for certain fit indices, most importantly RMSEA. Given these results, and the fact that the bifactor model allowed us to directly evaluate the usefulness of the subscales and the total scale in one combined model, the bifactor model was selected for further analyses. In addition, the modified correlated traits model that demonstrated an adequate fit excluded the Awareness items and the model thus provided no information regarding the behavior of the Awareness items in relation to the other 30 items of the DERS.

### Reliability and Explained Common Variance

Omega and OmegaH values obtained from the estimated final selected bifactor model are presented in Table 4. The overall reliability values (Omega), which can be thought of as weighted coefficient alpha, of the total and subscales were all high (range .796–.963). The OmegaH value for the total scale was likewise high, indicating that almost all of the reliable variance in total scores ( $.852/.962 = .885$ ) could be attributed to the general factor, assumed to measure general emotion dysregulation. Despite high omega values for all subscales, low

OmegaH values for the Goals, Impulse and Strategies subscales, however, indicate that most of the reliable variance on these subscales was accounted for by the reliable variance due to the general factor. The opposite was true for the Awareness and Clarity subscales that both had relatively high OmegaH and subsequently explained more of the reliable variance relative to the general factor (proportion = .94 and .70 respectively).

The estimate of ECV statistics was .86. This suggested that most of the common variance was accounted for by the general factor and thus only a small proportion of the variance ( $1 - .86 = .14$ ) appeared to be explained by the subscale factors beyond the general factor.

### Measurement and Structural Invariance across Anorexia Nervosa and Bulimia Nervosa

**Measurement Invariance** The bifactor solution fitted adequately in both the anorexia nervosa and bulimia nervosa subsamples (see Table 3). The combined multigroup configural model that posed no equality restrictions on intercepts and loadings across groups also fitted adequately (see Table 3). Multigroup CFA analyses revealed that a model that constrained intercepts and loadings (i.e., strong measurement invariance) to be equal across eating disorder groups did not significantly degrade model fit relative to the configural model,  $\Delta\chi^2(94) = 101.216, p = .287$  and  $\Delta CFI < .001$ . This suggested that strong measurement invariance assumption was met, and we could proceed and test for differences in structural parameters across groups (i.e., factor variance and factor means).

**Structural Invariance** Multigroup CFA analyses revealed that a model that constrained factor variance to be equal across groups did not significantly degrade model fit relative to a model in which factor variance were set as free parameters in each group (assuming strong measurement invariance in line with previous findings),  $\Delta\chi^2(8) = 6.382, p = .604$  and

<sup>2</sup> Given that the scaled chi square test for MLR produced a negative value due to a negative correction factor and thus could not be used, we instead used the z-value for the specific test of the covariance between Awareness and Clarity factor to obtain the chi-two value for model comparison with one degree of freedom (as two times the z-value approximately follow the chi-square distribution).



**Table 3** Estimated confirmatory factor analysis models for DERS-36 using the MLR estimator

Model		Chi-2	df	CFI	SRMR	RMSEA	[90% CI]
Unidimensional	Model 1.1	7320.013	594	0.592	0.111	0.115	[0.113, 0.117]
	Model 1.2	5164.097	405	0.663	0.092	0.117	[0.114, 0.120]
Trait	Model 2.1	2440.130	579	0.887	0.065	0.061	[0.059, 0.064]
	Model 2.2	1729.675	395	0.905	0.058	0.063	[0.060, 0.066]
Higher-order	Model 3.1	2756.743	588	0.869	0.082	0.066	[0.063, 0.068]
	Model 3.2	2472.762	587	0.886	0.067	0.061	[0.059, 0.064]
Bivariate	Model 4.1	2382.028	558	0.889	0.073	0.062	[0.059, 0.064]
	<b>Model 4.2</b>	<b>2010.668</b>	<b>557</b>	<b>0.912</b>	<b>0.049</b>	<b>0.055</b>	<b>[0.053, 0.058]</b>
Measurement invariance							
Bifactor	AN	1047.596	557	0.914	0.059	0.057	[0.052, 0.062]
	BN	1153.135	557	0.906	0.051	0.058	[0.053, 0.063]
Configural		2199.933	1114	0.910	0.055	0.057	[0.054, 0.061]
Scalar		2293.930	1208	0.910	0.060	0.055	[0.052, 0.059]

Best fitted model in bold

CFI, Comparative Fit Index; SRMR, Standardized Root Mean Residual; RMSEA, Root Mean-square error of approximation; AN, Anorexia Nervosa; BN, Bulimia Nervosa; CI, Confidence interval

$\Delta$ CFI <.001. This indicated that dispersion in factor scores were similar across groups.

Multigroup CFA analysis revealed that the bifactor model which held intercepts, loadings, factor variance and factor means equal across groups, fitted worse than an identical model but in which factor means were free parameters in bulimia nervous subsamples (using the anorexia nervosa diagnosis as the reference group with a mean of zero),  $\Delta\chi^2(7) = 22.284$ ,  $p = .002$ . When inspecting specific mean differences, a statistically significant difference was detected for the factor mean of the general factor ( $z = 2.163$ ,  $p = .031$ ,  $d = .214$ ) with participants with bulimia nervosa having a higher factor mean than those with anorexia nervosa (with a factor mean difference of .099 points). There was also statistically significant difference on the Clarity factor with participants with bulimia nervosa having a lower factor mean (mean difference = .136) than those with anorexia nervosa ( $z = 2.298$ ,  $p = .022$ ,  $d = .237$ ). There was also a tendency that participants with bulimia nervosa had a lower mean on the Strategies factor (mean difference = .069) and on the Nonacceptance factor (mean difference = .155) than those with anorexia nervosa, but these differences only approached and did not reach conventional level of statistical significance ( $z = 1.918$ ,  $p = 0.055$ ,  $d = .277$ ,  $z = 1.703$ ,  $p = .089$ ,  $d = .167$ , respectively). None of the other comparisons were statistically significant (all  $p$ 's > .6 and  $d$ 's < .05).

### Factor Structure of the DERS-16

The bifactor model that was established as the best-fitted model for all the 36 items was used to evaluate the factor structure of the 16-item version of the DERS in the full sample. Given that this shortened version did not include the Awareness

scale, no covariance could subsequently be added between this subfactor and the Clarity factor (as model 4.2 estimated in the full sample). In addition, due to the fact that the Clarity subscale only had two indicators, loadings were constrained to 1 for these items to make the model identified globally. The bifactor model provided a good fit to the data ( $\chi^2(89) = 379.985$ ,  $p < .001$ ; RMSEA = .062 [95% CI .055, .068]; CFI = .954; SRMR = .04).

### Criterion Validity and Diverging Associations with Eating Pathology

To examine the associations between dimensions of emotion dysregulation (i.e., latent factors) and eating pathology and whether the subfactors could account for unique variance over and above the general factor, we incorporated EDE-Q (as a measure of eating pathology) as an observed dependent variable into the best-fitted bifactor model. First, a model was run in which the EDE-Q was regressed on the general factor only (and all the subfactors effects on EDE-Q were constrained to zero.) Second, a model was run in which EDE-Q was regressed on all the subfactors and the general factor. The contribution of all subfactors ability to account for unique variance in the outcome above the contribution of the general factor was examined by improvement in global model fit between models using the scaled chi-square difference test. This analytic procedure was repeated for individuals with bulimia nervosa and individuals with anorexia nervosa.

For the anorexia nervosa subsample, the general factor was statistically significantly associated with scores on the EDE-Q ( $z = 5.099$ ,  $p < .001$ ) and accounted for approximately 15% of the variance ( $R^2 = .154$ ). The subfactors accounted for an

additional 2% of the variance, but the overall contribution was not statistically significant,  $\Delta\chi^2(6) = 9.509$ ,  $p = .147$ . For the bulimia nervosa subsample, the general factor was statistically significantly associated with scores on the EDE-Q ( $z = 5.245$ ,  $p < .001$ ) and accounted for approximately 14% of the variance ( $R^2 = .144$ ). The subsfactors accounted for an additional 5.7% of the variance, a contribution that was statistically significant,  $\Delta\chi^2(6) = 13.055$ ,  $p = .042$ . However, despite accounting for a significant contribution together, none of the estimates of the effects of subsfactors on EDE-Q had a statistically significant individual contribution in the model (all  $p$ 's  $> .06$ ; all standardized beta's  $< .17$ ).

## Discussion

### Factor Structure of DERS

The factor structure of the DERS-36 in a large eating disorder sample was examined using CFA with eight different models. Results indicated that a modified bifactor model that allowed the subscale Awareness to correlate with the subscale Clarity (model 4.2) provided best fit. This model was the only model to reach acceptable or good fit on all indices. Further, it was the only model to reach good fit according to our primary measure of fit (i.e., RMSEA). A similar bifactor solution was also found to be a good fit for the DERS-16. The bifactor model suggest that the composite score has more than one source of variance, meaning that the scores of all items on the DERS-36 and DERS-16 are influenced by more than one distinct underlying construct (i.e. subscales) beyond the contributions from the general factor. A bifactor model of the DERS-36 and DERS-16 has previously been suggested in two studies (Hallion et al. 2018; Osborne et al. 2017). Similar to our model, both of these studies made modifications to the model due to identified problems with the Awareness subscale. Osborne et al. (2017) tested a bifactor model that excluded items of Awareness scale from loading onto the total scale but permitted a correlation between this subfactor and the Clarity factor. Hallion et al. (2018) tested a bifactor model with all items from the Awareness subscale excluded. Thus, while the overall structure differed slightly with previous tested bifactor models, our result aligns with previous examinations of the factor structure of DERS in other clinical populations.

In comparison with the other tested bifactor models in other clinical samples, our bifactor model also allowed us to explicitly test the contribution of the Awareness subscale in accounting for unique and shared variance. The Awareness concept or similar constructs are common features in several emotion regulation models as well as in several models of psychiatric pathology (D'Agostino et al. 2017) and specific models of eating disorder pathology (Lavender et al. 2015). Keeping the Awareness subscale is also in line with recommendations

from the study by Fowler et al. (2014). While their study did not test a bifactor solution, the authors concluded that removal of the Awareness subscale did nothing to improve fit in their sample of 592 adult patients with severe mental illness. The only previous study of the factor structure of the DERS-36 in an eating disorder sample also found acceptable fit for a (correlated-traits) solution that included Awareness (Wolz et al. 2015). As for the DERS-16, the results from the present study indicated a good fit for a bifactor model, a finding that is in line with the results from Hallion et al. (2018) that is the only previous study to investigate the factor structure of the DERS-16 in a psychiatric clinical sample. While general recommendations on the factor structure are difficult to provide given different populations and methods in these studies, the present study adds additional support to this factor structure in another clinical population.

### Reliability and Explained Common Variance

Reliability of DERS-36 for individuals with eating disorder was examined in terms of Omega (comparable to a weighted coefficient alpha), OmegaH (reliable variance accounted for by the particular scale excluding reliable variance accounted for by the general factor), and ECV (the degree of common variance explained by the general factor). Results showed that OmegaH and the explained proportion of variance were high for DERS total, indicating that most of the explained common variance can be attributed to the total scale. Regarding the subscales, results varied. The results from the present study as well as the study by Osborne et al. (2017) are in line with earlier findings suggesting that the large majority of psychological multidimensional measures show similar tendencies with high OmegaH and ECV for the measures total scale and corresponding low OmegaH for subscales (Rodriguez et al. 2016a). There is no consensus in the field regarding specific benchmarks for evaluating OmegaH, but values greater than .50 or .75 have been suggested (Reise et al. 2013). Comparing with these cutoffs, our results indicate that half of the DERS subscales (Awareness, Clarity, and NonAcceptance) reach sufficient or close to sufficient OmegaH values and the other half present lower values. In this context it is important to note that all Omega values (comparable to a weighted coefficient alpha) were high for all subscales.

The subscales Goals, Impulse, and Strategies all exhibited low OmegaH, indicating that these subscales contribute with little unique explanatory value beyond the contributions of the total scale. The Awareness, Clarity and Nonacceptance subscales showed medium to relatively high OmegaH-values. Only one previous study by Osborne et al. (2017) has presented OmegaH for the DERS. Results were similar with low OmegaH for Goals, Impulse and Strategies as well as high OmegaH for the subscale Clarity. Osborne et al. (2017) did

**Table 4** Omega and Omega Hierarchical for total and subscales of the DERS-36

Scale	Omega	OmegaH	Proportion of reliable variance (OmegaH/Omega)
Total	0.963	0.852	0.885
Non-Accept	0.897	0.479	0.534
Goals	0.895	0.249	0.278
Impulse	0.918	0.261	0.284
Awareness	0.796	0.749	0.941
Strategies	0.904	0.098	0.108
Clarity	0.842	0.589	0.700

OmegaH, Omega Hierarchical

not provide OmegaH for the Awareness subscale as their modified bifactor model prevented Awareness items from loading onto the total scale. OmegaH for DERS total were similar to ours.

There is no consensus in the literature regarding the use of the DERS total score or the use of subscale scores in clinical populations. For example, Fowler et al. (2014) cautions against the use of DERS total, since the results from the higher-order models in their study showed a weak fit, suggesting that future use should focus primarily on the six subscales, whereas Osborne et al. (2017) recommended the use of a general or total score from the modified DERS with only five subscales (excluding the Awareness subscale). Taken together, the results from the present study indicate that there may be reasons for being cautious when interpreting scores of some of the DERS subscales in an eating disorder population as most of the reliable variance might be accounted for by a general factor representing shared variance among all the items. This is especially true for Strategies. Regarding Awareness, the high OmegaH values indicate that the subscale contributes a high proportion of unique variance over and above the variance explained by the general factor. This could lead to questions regarding the underlying latent construct, as it is possible that Awareness measures some different aspect of emotion regulation than the rest of the items of the DERS (Bardeen et al. 2012; Lee et al. 2016). The present findings dovetail with previous studies that have found small inter-correlations between the Awareness factor and the other factors (Fowler et al. 2014; Osborne et al. 2017; Perez et al. 2012; Wolz et al. 2015). On the other hand, the Awareness subscale is the only one to meet the preferred benchmark of .75 or higher OmegaH suggested by Reise et al. (2013). It might thus be possible to use this scale on its own, but more difficult to know whether the scale actually taps into to a more general emotion regulation ability as measured by the DERS.

### Measurement and Structural Invariance across Anorexia Nervosa and Bulimia Nervosa

The present study is the first to test for measurement and structural invariance for the DERS-36 across the eating

disorder subgroups anorexia nervosa and bulimia nervosa. Results showed that strong measurement invariance assumption was met. This suggests that the DERS-36 measures the same underlying latent construct in both groups, meaning that it is possible to make relevant comparisons between patients with anorexia nervosa and bulimia nervosa in terms of emotion regulation difficulties as measured by the DERS-36. This result is of importance in the field as comparisons and correlations between emotion regulation difficulties measured by the DERS in different eating disorder subgroups is a common study aim in eating disorder research (e.g. Brockmeyer et al. 2014; Lavender et al. 2015; Segal and Golan 2016; Svaldi et al. 2012). Further analyses showed that the dispersion in factor scores was similar across groups. Regarding mean factor score differences between anorexia nervosa and bulimia nervosa the results indicated small but statistically significant differences for DERS total as well as Clarity, with higher total mean and lower mean on Clarity in the bulimia nervosa group. No other comparisons were statistically significant. Taken together, patients with anorexia nervosa and bulimia nervosa seem to have fairly similar emotion regulation difficulties as measured by the DERS-36, a finding that is consistent with previous research (e.g., Brockmeyer et al. 2014; Lavender et al. 2015; Monell et al. 2018; Svaldi et al. 2012).

### Criterion Validity

Results from bifactor models revealed that the general factor of emotion dysregulation as measured with the DERS in the best-fitted bifactor model was strongly associated with eating psychopathology in both anorexia nervosa and bulimia nervosa subsamples (15% and 14% respectively). These findings are in line with previous research that has shown that emotion dysregulation is an important feature of eating pathology across eating disorders (e.g., Brockmeyer et al. 2014; Lavender et al. 2015). Interestingly, once the variance by the general factor was accounted for in the model, the overall contribution of the subfactors was relatively small in both diagnostic groups (2% and 5%) and was not statistically significant in the anorexia nervosa subsample. None of the

subfactors had a statistically significant contribution on their own in accounting for eating pathology. By incorporating an external criterion of relevance for eating disorders (eating disorder psychopathology), these findings corroborate our other results showing that most of the reliable variance was accounted for by the general factor. Of course, one might argue that it is difficult to know what it actually means to be high (or low) on a subscale that is supposed to tap into a specific emotion regulation skill when parcelling out general emotion ability. Indeed, concerns in terms of the interpretation of a bifactor model have been raised (Bonifay et al. 2017; Murray and Johnson 2013). On the other hand, if most of the reliable variance is indeed accounted for by a general factor, and in the absence of evidence for reliability of the subscale scores, it is difficult to provide a meaningful interpretation of the associations between observed subscales scores of the DERS and other important variables. At this time point, it is difficult to make definitive conclusions as more research on the validity of the DERS in individuals with eating disorders is warranted. Still, in this subsample of individuals with eating disorders, our results highlight that the subfactors cannot be considered to reflect broad, independent abilities because they include a strong contribution of the general factor.

### Strengths and Limitations

The findings of this study should be interpreted considering some strengths and limitations. First, a major strength is the large psychiatric eating disorder sample that provides an ecologically valid context for assessment. The DERS was developed for populations with clinical problems regarding emotion regulation and different kinds of psychopathology (Gratz and Roemer 2004). The present study makes important contributions regarding the psychometric properties and use of the DERS-36 and DERS-16 in the field of clinical psychiatric eating disorders. Second, the present study investigates the latent structure of the DERS with several models, including a bifactor model that has previously been suggested to be suitable for measures that are assumed to have a multidimensional structure (such as the DERS). The use of a bifactor model also allowed for investigations of reliability in terms of Omega, OmegaH, and explained common variance, giving important information regarding the previously identified problems with certain subscales. To our knowledge, there is only one previous study that has reported OmegaH for the DERS and the present study is the first to present OmegaH for all six subscales. Lastly, the test for measurement invariance is an important strength as this provides essential information regarding comparisons of emotion regulation across eating disorder diagnostic subgroups.

There are also a few limitations with the current study. The DERS was added as an optional measure in the Stepwise register and criteria for clinics' decisions in inclusion are not

recorded. A previous study using the DERS from the Stepwise register ( $N=999$ , only women) could not find any differences between patients with or without DERS registration regarding variables such as age, body mass index, eating disorder characteristics and psychiatric comorbidity (see Monell et al. 2018). The fact that the present study largely shares this sample makes it plausible that this should be true for this study as well. Another related aspect concerns the nesting of individuals within clinics that could potentially have influenced the results. It is however important to note that sensitivity analyses were run that controlled for the nesting of observations within units and results were not altered materially from those presented.

Second, although the bifactor model comes with certain advantages, some concerns have also been raised. The bifactor model has shown tendencies to outperform other models in terms of fit statistics, possibly due to statistical bias, making model comparisons and interpretation of fit indices more challenging (Bonifay et al. 2017; Murray and Johnson 2013). It is also important to note that the differences in fit between some of the factor models were not substantial and even though the bifactor model provided an adequate fit to the data in the current sample, we cannot know for sure if this is a valid representation of the structure in the population. Additionally, conceptual concerns have also been raised regarding fitting bifactor models to psychological constructs, for example related to interpretation of the meaning of a bifactor structure (Bonifay et al. 2017; Murray and Johnson 2013). Notwithstanding these concerns, as pointed out by Bonifay et al. (2017) and others, the bifactor model still might be of importance for the development and evaluation of measurements given that it offers a unique opportunity to calculate various useful indices (e.g., OmegaH, EVC) that can assist in determining whether a measure has an acceptable true score variance and the extent to which subscales scores are reliable after accounting for the general factor.

Our study did not examine validity by comparing the DERS with other established emotion regulation measures. It should, however, be noted that the validity of the DERS has previously been examined by associations with other relevant self-assessed variables in both nonclinical and clinical populations, including eating disorders (e.g., Brockmeyer et al. 2014; Monell et al. 2015; Svaldi et al. 2012). It would be most informative if future research also makes efforts to establish the validity of the DERS in other ways than merely studying correlations among self-report assessments, for example, using data from multimethod and longitudinal studies or experiments to assess the evidence of criterion relevance, discriminative validity and responsiveness to treatment among individuals with different eating disorders.

The sample in the current study was limited to treatment seeking adults (mostly females) with eating disorder without any reference group (e.g., normal controls). This, naturally,



makes it difficult to determine the extent to which these results can be generalized to other clinical and nonclinical populations. Also, regarding measurement and structural invariance for diagnostic subgroups, the present study is limited to comparisons between individuals with anorexia nervosa and bulimia nervosa. For future research, the final bifactor solution needs replication as this is the first study to test that specific model in a clinical sample of eating disorder. There is also a need to broaden the investigation between subgroups to different kinds of eating disorders as well as comparisons between subgroups defined in other ways than by the diagnostic manual. The eating disorder diagnostic criteria have been criticized (e.g., Forbush and Wildes 2017) due to the overlap of behavioral symptoms and the fact that individuals with eating disorders tend to move between different sub-diagnoses over time. The findings of this study also warrant replication in a sample of younger patients.

## Conclusions

The present study is the first to examine the factor structure of the DERS-36 and DERS-16 in a large clinical eating disorder sample, and to test a bifactor model as well as measurement invariance across diagnostic subgroups in this population. Results indicated good fit for a bifactor model with permitted correlations between the subscales Awareness and Clarity. Analyses of reliability suggest that the DERS-36 and the DERS-16 total scale are reliable, while results from the reliability analyses of the subscales varied. The interpretation of some specific subscales is less clear as most of the reliable variance accounted for by these subscales overlapped with the variance accounted for by the general factor. The general factor also accounted for most of the variance in eating pathology and none of the subfactors had a statistically significant contribution on their own. Finally, results also showed that criteria for strong measurement invariance were met, indicating that meaningful comparisons of the eating disorder subgroups (i.e., anorexia nervosa and bulimia nervosa) using the DERS are possible. Taken together, findings indicate that the use of the total scale can be recommended when administered in an eating disorder population.

**Acknowledgements** Open access funding provided by Linköping University.

## Compliance with Ethical Standards

**Conflict of Interest** Line Nordgren, Elin Monell, Andreas Birgegård, Johan Bjureberg and Hugo Hesser declare that they have no conflict of interest.

**Experiment Participants** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964

Helsinki declaration and its later amendments or comparable ethical standards.

**Informed Consent** Informed consent was obtained from all individual participants included in the study.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5™* (5th ed.). Arlington, VA, US: American Psychiatric Publishing, Inc..
- Bagby, R. M., Taylor, G. J., & Parker, J. D. A. (1994). The twenty-item Toronto alexithymia scale: II. Convergent, discriminant, and concurrent validity. *Journal of Psychosomatic Research*, 38(1), 33–40. [https://doi.org/10.1016/0022-3999\(94\)90006-X](https://doi.org/10.1016/0022-3999(94)90006-X).
- Bardeen, J. R., Fergus, T. A., & Orcutt, H. K. (2012). An examination of the latent structure of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, 34(3), 382–392. <https://doi.org/10.1007/s10862-012-9280-y>.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>.
- Birgegård, A., Björck, C., & Clinton, D. (2010). Quality assurance of specialised treatment of eating disorders using large-scale internet-based collection systems: Methods, results and lessons learned from designing the Stepwise database. *European Eating Disorders Review*, 18(4), 251–259. <https://doi.org/10.1002/erv.1003>.
- Bjureberg, J., Ljótsson, B., Tull, M. T., Hedman, E., Sahlin, H., Lundh, L. G., ... Gratz, K. L. (2016). Development and validation of a brief version of the difficulties in emotion regulation scale: The DERS-16. *Journal of Psychopathology and Behavioral Assessment*, 38(2), 284–296. <https://doi.org/10.1007/s10862-015-9514-x>.
- Bonifay, W., Lane, S. P., & Reise, S. P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science*, 5(1), 184–186. <https://doi.org/10.1177/2167702616657069>.
- Bostan, C. M., & Zaharia, D. V. (2016). Emotional dysregulation - factor structure and consistency in the Romanian version of the difficulties in emotion regulation scale (DERS). *Annals of the Al. I. Cuza University, Psychology Series*, 25(2), 57–77.
- Brockmeyer, T., Skunde, M., Wu, M., Bresslein, E., Rudolfsky, G., Herzog, W., & Friederich, H. C. (2014). Difficulties in emotion regulation across the spectrum of eating disorders. *Comprehensive Psychiatry*, 55(3), 565–571. <https://doi.org/10.1016/j.comppsy.2013.12.001>.
- Burns, E. E., Fischer, S., Jackson, J. L., & Harding, H. G. (2012). Deficits in emotion regulation mediate the relationship between childhood abuse and later eating disorder symptoms. *Child Abuse and Neglect*, 36(1), 32–39. <https://doi.org/10.1016/j.chiabu.2011.08.005>.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation*

- Modeling*, 9(2), 233–255. [https://doi.org/10.1207/S15328007SEM0902\\_5](https://doi.org/10.1207/S15328007SEM0902_5).
- Cooper, Z., & Fairburn, C. (1987). The eating disorder examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders*, 6(1), 1–8. [https://doi.org/10.1002/1098-108X\(198701\)6:1<1::AID-EAT2260060102>3.0.CO;2-9](https://doi.org/10.1002/1098-108X(198701)6:1<1::AID-EAT2260060102>3.0.CO;2-9).
- 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*, 47(6), 630–639. <https://doi.org/10.1002/eat.22278>.
- D'Agostino, A., Covanti, S., Rossi Monti, M., & Starcevic, V. (2017). Reconsidering emotion dysregulation. *Psychiatric Quarterly*, 88(4), 807–825. <https://doi.org/10.1007/s11126-017-9499-6>.
- de Man Lapidoth, J., & Birgegård, A. (2010). *Validation of the structured eating disorder interview (SEDI) against the eating disorder examination (EDE)*. Stockholm: Karolinska institutet.
- Fairburn, C. G., & Bèglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16(4), 363–370.
- First, M., Spitzer, R. L., Gibbon, M. L., & Williams, J. (2002). Structured clinical interview for DSM-IV-TR Axis I disorders, research version, non-patient edition. (SCID-I/P). New York: Biometrics research, New York state psychiatric institute.
- Forbush, K. T., & Wildes, J. E. (2017). Application of structural equation mixture modeling to characterize the latent structure of eating pathology. *International Journal of Eating Disorders*, 50(5), 542–550. <https://doi.org/10.1002/eat.22634>.
- Fowler, J. C., Charak, R., Elhai, J. D., Allen, J. G., Frueh, B. C., & Oldham, J. M. (2014). Construct validity and factor structure of the difficulties in emotion regulation scale among adults with severe mental illness. *Journal of Psychiatric Research*, 58, 175–180. <https://doi.org/10.1016/j.jpsychires.2014.07.029>.
- Garke, M., Sörman, K., Jayaram-Lindström, N., Hellner, C., & Birgegård, A. (2019). Symptom shifting and associations with mental illness: A transdiagnostic approach applied to eating disorders. *Journal of Abnormal Psychology*, 128(6), 585–595. <https://doi.org/10.1037/abn0000425>.
- Garner, D. M., Olmstead, M. P., & Polivy, J. (1983). Multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders*, 2(2), 15–34.
- Ghorbani, F., Khosravani, V., Shari, F., & Jamaati, R. (2017). The alexithymia, emotion regulation, emotion regulation difficulties, positive and negative affects, and suicidal risk in alcohol-dependent outpatients. *Psychiatry Research*, 252, 223–230. <https://doi.org/10.1016/j.psychres.2017.03.005>.
- Gómez-Expósito, A., Wolz, I., Fagundo, A. B., Granero, R., Steward, T., Jiménez-Murcia, S., et al. (2016). Correlates of non-suicidal self-injury and suicide attempts in bulimic spectrum disorders. *Frontiers in Psychology*, 7, 1244. <https://doi.org/10.3389/fpsyg.2016.01244>.
- Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement*, 66(6), 930–944. <https://doi.org/10.1177/0013164406288165>.
- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation. *Journal of Psychopathology and Behavioral Assessment*, 26(1), 41–54. <https://doi.org/10.1023/B:JOBA.0000007455.08539.94>.
- Gratz, K. L., Rosenthal, M. Z., Tull, M. T., Lejuez, C. W., & Gunderson, J. G. (2006). An experimental investigation of emotion dysregulation in borderline personality disorder. *Journal of Abnormal Psychology*, 115(4), 850–855. <https://doi.org/10.1037/0021-843X.115.4.850>.
- Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, 26(1), 1–26. <https://doi.org/10.1080/1047840X.2014.940781>.
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, 85(2), 348–362. <https://doi.org/10.1037/0022-3514.85.2.348>.
- Hallion, L. S., Steinman, S. A., Tolin, D. F., & Diefenbach, G. J. (2018). Psychometric properties of the difficulties in emotion regulation scale (DERS) and its short forms in adults with emotional disorders. *Frontiers in Psychology*, 9, 539. <https://doi.org/10.3389/fpsyg.2018.00539>.
- Haynos, A. F., Roberto, C. A., & Attia, E. (2015). Examining the associations between emotion regulation difficulties, anxiety, and eating disorder severity among inpatients with anorexia nervosa. *Comprehensive Psychiatry*, 60, 93–98. <https://doi.org/10.1016/j.comppsy.2015.03.004>.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>.
- Kököneyi, G., Urbán, R., Reinhardt, M., Anna, J., & Demetrovics, Z. (2014). The difficulties in emotion regulation scale: Factor structure in chronic pain patients. *Journal of Clinical Psychology*, 70(6), 589–600. <https://doi.org/10.1002/jclp.22036>.
- Lavender, J. M., Wonderlich, S. A., Engel, S. G., Gordon, K. H., Kaye, W. H., & Mitchell, J. E. (2015). Dimensions of emotion dysregulation in anorexia nervosa and bulimia nervosa: A conceptual review of the empirical literature. *Clinical Psychology Review*, 40, 111–122. <https://doi.org/10.1016/j.cpr.2015.05.010>.
- Lee, D. J., Witte, T. K., Bardeen, J. R., Davis, M. T., & Weathers, F. W. (2016). A factor analytic evaluation of the difficulties in emotion regulation scale. *Journal of Clinical Psychology*, 72(9), 933–946. <https://doi.org/10.1002/jclp.22297>.
- Lindvall Dahlgren, C., & Wisting, L. (2016). Transitioning from DSM-IV to DSM-5: A systematic review of eating disorder prevalence assessment. *International Journal of Eating Disorders*, 49(11), 975–997. <https://doi.org/10.1002/eat.22596>.
- Luce, K. H., & Crowther, J. H. (1999). The reliability of the eating disorder examination—Self-report questionnaire version (EDE-Q). *International Journal of Eating Disorders*, 25(3), 349–351. [https://doi.org/10.1002/\(SICI\)1098-108X\(199904\)25:3<349::AID-EAT15>3.0.CO;2-M](https://doi.org/10.1002/(SICI)1098-108X(199904)25:3<349::AID-EAT15>3.0.CO;2-M).
- Mallorquí-Bagué, N., Vintró-Alcaraz, C., Sánchez, I., Riesco, N., Agüera, Z., Granero, R., ... Fernández-Aranda, F. (2018). Emotion regulation as a transdiagnostic feature among eating disorders: Cross-sectional and longitudinal approach. *European Eating Disorders Review*, 26(1), 53–61. <https://doi.org/10.1002/erv.2570>.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543. <https://doi.org/10.1007/BF02294825>.
- Miguel, F. K., Giromini, L., Colombaroli, M. S., Zuanazzi, A. C., & Zennaro, A. (2017). A Brazilian investigation of the 36- and 16-item difficulties in emotion regulation scales. *Journal of Clinical Psychology*, 73(9), 1146–1159. <https://doi.org/10.1002/jclp.22404>.
- Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. V. (2004). Validity of the eating disorder examination questionnaire (EDE-Q) in screening for eating disorders in community samples. *Behaviour Research and Therapy*, 42(5), 551–567. [https://doi.org/10.1016/S0005-7967\(03\)00161-X](https://doi.org/10.1016/S0005-7967(03)00161-X).
- Monell, E., Högdahl, L., Mantilla, E. F., & Birgegård, A. (2015). Emotion dysregulation, self-image and eating disorder symptoms in

- university women. *Journal of Eating Disorders*, 3, 44. <https://doi.org/10.1186/s40337-015-0083-x>.
- Monell, E., Clinton, D., & Birgegård, A. (2018). Emotion dysregulation and eating disorders—associations with diagnostic presentation and key symptoms. *International Journal of Eating Disorders*, 51, 921–930. <https://doi.org/10.1002/eat.22925>.
- Moskowitz, L., & Weiselberg, E. (2017). Anorexia Nervosa/Atypical Anorexia Nervosa. *Current Problems in Pediatric and Adolescent Health Care*, 47(4), 70–84. <https://doi.org/10.1016/j.cppeds.2017.02.003>.
- Murray, A. L., & Johnson, W. (2013). The limitations of model fit in comparing the bi-factor versus higher-order models of human cognitive ability structure. *Intelligence*, 41(5), 407–422. <https://doi.org/10.1016/j.intell.2013.06.004>.
- Muthén, B. O. (2015). *Mplus user's guide* (Seventh ed.). Los Angeles: Muthén & Muthén.
- Neumann, A., van Lier, P., Gratz, K. L., & Koot, H. (2010). Multidimensional assessment of emotion regulation difficulties in adolescents using the difficulties in emotion regulation scale. *Assessment*, 17(1), 138–149.
- Osborne, T. L., Michonski, J., Sayers, J., Welch, S. S., & Anderson, L. K. (2017). Factor structure of the difficulties in emotion regulation scale (DERS) in adult outpatients receiving dialectical behavior therapy (DBT). *Journal of Psychopathology and Behavioral Assessment*, 39(2), 355–371. <https://doi.org/10.1007/s10862-017-9586-x>.
- Perez, J., Venta, A., Garnaat, S., & Sharp, C. (2012). The difficulties in emotion regulation scale: Factor structure and association with non-suicidal self-injury in adolescent inpatients. *Journal of Psychopathology and Behavioral Assessment*, 34(3), 393–404. <https://doi.org/10.1007/s10862-012-9292-7>.
- Pisetsky, E. M., Thornton, L. M., Lichtenstein, P., Pedersen, N. L., & Bulik, C. M. (2013). Suicide attempts in women with eating disorders. *Journal of Abnormal Psychology*, 122(4), 1042–1056. <https://doi.org/10.1037/a0034902>.
- Racine, S. E., & Wildes, J. E. (2015). Dynamic longitudinal relations between emotion regulation difficulties and anorexia nervosa symptoms over the year following intensive treatment. *Journal of Consulting and Clinical Psychology*, 83(4), 785–795. <https://doi.org/10.1037/ccp0000011>.
- Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, 32(4), 329–353. [https://doi.org/10.1207/s15327906mbr3204\\_2](https://doi.org/10.1207/s15327906mbr3204_2).
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667–696. <https://doi.org/10.1080/00273171.2012.715555>.
- Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, 95(2), 129–140. <https://doi.org/10.1080/00223891.2012.725437>.
- Rieffe, C., Terwogt, M. M., Petrides, K. V., Cowan, R., Miers, A. C., & Tolland, A. (2007). Psychometric properties of the emotion awareness questionnaire for children. *Personality and Individual Differences*, 43, 95–105.
- Ritschel, L. A., Tone, E. B., Schoemann, A. M., & Lim, N. E. (2015). Psychometric properties of the difficulties in emotion regulation scale across demographic groups. *Psychological Assessment*, 27(3), 944–954. <https://doi.org/10.1037/pas0000099>.
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016a). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, 98(3), 223–237. <https://doi.org/10.1080/00223891.2015.1089249>.
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016b). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, 21(2), 137–150. <https://doi.org/10.1037/met0000045>.
- Sartaş-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(1), 12–19. <https://doi.org/10.1027/1015-5759/a000199>.
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4), 507–514. <https://doi.org/10.1007/BF02296192>.
- Sawyer, S. M., Whitelaw, M., Le Grange, D., Yeo, M., & Hughes, E. K. (2016). Physical and psychological morbidity in adolescents with atypical anorexia nervosa. *Pediatrics*, 137(4), e20154080. <https://doi.org/10.1542/peds.2015-4080>.
- Segal, A., & Golan, M. (2016). Differences in emotion regulation along the eating disorder spectrum: Cross sectional study in adolescents out patient care. *Journal of Psychology & Clinical Psychiatry*, 6(1), 00314. <https://doi.org/10.15406/jcpsy.2016.06.00314>.
- Shahabi, M., Hasani, J., & Bjureberg, J. (2018). Psychometric properties of the brief Persian version of the difficulties in emotion regulation scale (the DERS-16). *Assessment for Effective Intervention*. <https://doi.org/10.1177/1534508418800210>.
- Sloan, E., Hall, K., Moulding, R., Bryce, S., Mildred, H., & Staiger, P. K. (2017). Emotion regulation as a transdiagnostic treatment construct across anxiety, depression, substance, eating and borderline personality disorders: A systematic review. *Clinical Psychology Review*, 57, 141–163. <https://doi.org/10.1016/j.cpr.2017.09.002>.
- Svaldi, J., Griepenstroh, J., Tuschen-Caffier, B., & Ehring, T. (2012). Emotion regulation deficits in eating disorders: A marker of eating pathology or general psychopathology? *Psychiatry Research*, 197(1–2), 103–111. <https://doi.org/10.1016/j.psychres.2011.11.009>.
- Tull, M. T., & Roemer, L. (2007). Emotion regulation difficulties associated with the experience of unmet panic attacks: Evidence of experiential avoidance, emotional nonacceptance, and decreased emotional clarity. *Behavior Therapy*, 38(4), 378–391. <https://doi.org/10.1016/j.beth.2006.10.006>.
- Velotti, P., Garofalo, C., Petrocchi, C., Cavallo, F., Popolo, R., & Dimaggio, G. (2016). Alexithymia, emotion dysregulation, impulsivity and aggression: A multiple mediation model. *Psychiatry Research*, 237, 296–303. <https://doi.org/10.1016/j.psychres.2016.01.025>.
- Venta, A., Hart, J., & Sharp, C. (2012). The relation between experiential avoidance, alexithymia and emotion regulation in inpatient adolescents. *Clinical Child Psychology and Psychiatry*, 18(3), 389–410. <https://doi.org/10.1177/1359104512455815>.
- Vieira, A. I., Ramalho, S., Brandão, I., Saraiva, J., & Gonçalves, S. (2016). Adversity, emotion regulation, and non-suicidal self-injury in eating disorders. *Eating Disorders*, 24(5), 440–452.
- Villalta, L., Smith, P., Hickin, N., & Stringaris, A. (2018). Emotion regulation difficulties in traumatized youth: A meta-analysis and conceptual review. *European Child & Adolescent Psychiatry*, 27, 527–544.
- Visted, E., Vøllestad, J., Nielsen, M. B., & Schanche, E. (2018). Emotion regulation in current and remitted depression: A systematic review and meta-analysis. *Frontiers in Psychology*, 9(756). <https://doi.org/10.3389/fpsyg.2018.00756>.
- Wang, J., & Wang, X. (2012). *Structural equation modeling: Applications using Mplus*. Chichester: John Wiley & Sons.
- Westerlund, M., & Santtila, P. (2018). A Finnish adaptation of the emotion regulation questionnaire (ERQ) and the difficulties in emotion regulation scale (DERS-16). *Nordic Psychology*, 70(4), 304–323. <https://doi.org/10.1080/19012276.2018.1443279>.
- Wolz, I., Agüera, Z., Granero, R., Jiménez-Murcia, S., Gratz, K. L., Menchón, J. M., & Fernández-Aranda, F. (2015). Emotion regulation in disordered eating: Psychometric properties of the difficulties in emotion regulation scale among Spanish adults and its

- interrelations with personality and clinical severity. *Frontiers in Psychology*, 6, 907. <https://doi.org/10.3389/fpsyg.2015.00907>.
- Yiugit, I., & Guzey Yiugit, M. (2017). Psychometric properties of Turkish version of difficulties in emotion rRgulation scale-brief form (DERS-16). *Current Psychology*, 7, 1–9. <https://doi.org/10.1007/s12144-017-9712-7>.
- Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's  $\alpha$ , Revelle's  $\beta$  and McDonald's  $\omega$ H: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, 70(1), 123–133. <https://doi.org/10.1007/s11336-003-0974-7>.

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