TPM

DEVELOPMENT OF THE DIFFICULTIES IN EMOTION REGULATION SCALE SHORT FORM (DERS-SF) WITH METAHEURISTIC ALGORITHMS

HAKAN KOĞAR
AKDENIZ UNIVERSITY, ANTALYA, TURKEY

The Difficulties in Emotion Regulation Scale (DERS) is one of the most widely used measures of emotion regulation. The general purpose of this research is to create the short form of the Difficulty in Emotion Regulation Scale (DERS-SF) by comparing the findings obtained from various metaheuristic algorithms (ACO, TS, and SA). The data of this study were collected from 398 undergraduate students. Correlated 6-factor and bifactor structures of DERS are tested with CFA. The model-data fit obtained from the correlated 6-factor structure created by the items selected by the ACO algorithms indicates a perfect fit. Also, hierarchical omega coefficients and standardized factor loadings show adequate psychometric properties. Strict measurement invariance was provided between gender groups. The correlations between the total score obtained from the Self-Compassion Scale and the nonacceptance, goals, impulse, strategies, awareness, and clarity factors of DERS-SF are negative and vary between moderate to high. As a result of this study, DERS-SF was developed as six factors and 18 items.

Keywords: DERS; DERS-SF; Metaheuristic algorithms; Short form selection; Confirmatory factor analysis.

Correspondence concerning this article should be addressed to Hakan Koğar, Department of Educational Sciences, Akdeniz University, Dumlupınar Bulvarı, 07058 Antalya, Turkey. Email: hkogar@gmail.com

Emotion regulation is broadly defined as the set of processes involved in coping with relatively strong emotional episodes (Kopp, 1989). It is an ability to monitor and alter emotional states in pursuit of goal-directed behavior (Gratz & Roemer, 2004). To achieve proper emotion regulation, people must be able to understand, accept, and modulate their emotions, as well as to adapt their behavior in response to various situations. Deficiencies in these competencies result in difficulties in achieving successful emotion regulation (Gratz & Roemer, 2004). With the development of efforts to measure emotion regulation, the effect of this psychological feature on other psychological features has also been examined. Emotion regulation failure may affect depression (Gross & Munoz, 1995), anxiety (Aldao et al., 2010), social difficulties due to poor emotional management (Dan-Glauser & Scherer, 2013), posttraumatic stress disorder (Ehring & Quack 2010; Tull et al., 2007), and borderline personality disorder (Gratz et al., 2006).

There are various measurement tools developed to measure emotion regulation. One of these tools, the Trait Meta Mood Scale (Salovey et al., 1995), can measure a part of emotion regulation. This measurement tool measures individual differences in a person's ability to reflect and manage their emotions (Salovey et al, 1995). Another measurement tool, the Negative Mood Regulation Scale (Catanzaro & Mearns, 1990), does not evaluate some important aspects of emotion regulation such as awareness of emotions, understanding of emotions, and acceptance of emotions. Based on all these shortcomings, Gratz and Roemer (2004) developed the Difficulties in Emotion Regulation Scale (DERS).

The DERS and Its Factor Structure

DERS is one of the most widely used measures of emotion regulation, with over 8,500 citations (Google scholar search in March 2022) attributed to the initial study. Gratz and Roemer (2004) found that a 6-factor correlated trait model ultimately best operationalized the original 36-item DERS. The factors included nonacceptance of emotional responses (*nonacceptance*), difficulties engaging in goal-directed activities (*goals*), impulse control difficulties (*impulse*), limited access to emotion regulation strategies (*strategies*), lack of emotional awareness (*awareness*), and lack of emotional clarity (*clarity*).

The DERS items are generally summed to create a total score and subscale scores. These total scores have been used to predict a variety of psychological phenomena, including posttraumatic stress (Tull et al., 2007). The original DERS was validated as a 6-factor multidimensional model. The multidimensional structure of the DERS has been examined in some studies (Dan-Glauser & Scherer, 2013; Neumann et al., 2010; Perez et al., 2012; Rugancı & Gençöz, 2010), and some results from the confirmatory factor analysis (CFA) have indicated adequate fit (Bardeen et al., 2012). In contrast to the multidimensional model, a bifactor modeling approach ensures which items of a measurement tool are loaded on both a general factor and a specific factor. As such, bifactor modeling can be used to determine whether a total score is sufficient to operationalize the construct assessed by a given measure (Reise, 2012). Osborne et al.'s (2017) study on the bifactor structure of DERS shows a good fit. In this study DERS' unidimensional, correlated factors, higher order, and bifactor structures were tested. The best fit indices were obtained from the bifactor model (RMSEA = .06; CFI = .95; TLI = .95). Unlike Osborne et al.'s (2017) study, Benfer et al. (2019) tested the unidimensional, 6-factor correlated, second-order and bifactor structure of DERS, but none of the examined models provided an adequate fit to the data. Therefore, they developed a modified version of the scale. Controversial issues related to the factor structure of such a scale provided the main motivation for this research.

Short Form Development

Short form development studies, which gained speed in the 1950s, focused primarily on measurement tools used for clinical evaluations as a result of the criticism against many items in intelligence and ability tests (Levy, 1968). Today, it has become an important interest, especially in psychology literature. Kruyen et al. (2013), in their study, which examined articles published in six important psychology journals over five years, stated that a total of 164 abbreviated tests were used in 170 articles. The motivation behind the development of the short form of only 70 of these tests was described. It was determined that only half of the short forms had construct validity proofs, and the reliability obtained in all but a few tests was lower than in the long form. Although this study shows that the use of short forms is very common, it demonstrates that there are many problems related to its development and use.

The interest in short form development is not limited to the field of psychology. It also significantly affects other areas such as education, as it provides financial savings, increases return to measurement tools, and saves time. When the literature is examined, it is possible to say that the methods used in short form development are divided into statistical and judgmental. It can be said that judgmental methods are handled within the framework of Classical Test Theory and are based on expert opinions. However, it is not appropriate to use only judgmental methods as it will enable the perceived fitness to be determined. Selecting the items with the highest contribution to reliability or selecting items with high factor loading or high interitem correlation, which are the most frequently used statistical methods, decides which items should be removed from the measurement tool, even if it is effective in maximizing the general validity or reliability of the measuring tool and determining that the last item number is insufficient (Kruyen et al., 2013). These methods are generally insufficient from a psychometric perspective because they are not interested in the factor structure of the measurement tool (Schroeders et al., 2016).

Algorithms Used in Short Form Development

Statistical methods based on simple item selection strategies are insufficient to select the correct short form item pool (Kleka & Soroko, 2018). The biggest advantage of metaheuristic algorithms over other statistical algorithms is that they approach short form development studies from a psychometric point of view. In this way, psychometric concepts such as structure validity and reliability are taken into account while developing the short form. In addition, the decision whether an item should be included in the specific factor is made by taking into account factors such as factor loadings (Byrne & Pachana, 2011; Noble et al., 2013). Creating a short form that has both sufficient internal structure and strong validity in terms of relationships with other variables is difficult with traditional short form development methods. Metaheuristic optimization algorithms (Dorigo et al., 2006) have the potential to solve these problems because they can simultaneously maximize more than one validity criterion for the developed short forms. Most of the metaheuristic algorithms are inspired by natural mechanisms such as the evolution of the foraging behavior of ants or honeybees. The use of these algorithms in the psychometric literature has been the development of automatic algorithms to generate short forms.

One of these algorithms, ant colony optimization (ACO), was inspired by the foraging behavior of ants. During their food search, ants leave pheromone tracks along their route, which will attract more ants to the route. At the start of the search, the ants will randomly search for the food source and return to the nest on the same route. Ants that find a shorter path to the food source will travel more frequently between the nest and the food source, thus accumulating more pheromones on their route. Pheromone levels on longer routes increase more slowly or evaporate over time. Higher pheromone levels will attract more ants, causing almost all ants to take the shortest path. This algorithm was developed by Colorni et al. (1991) as a solution to the traveling salesman problem. According to many studies, it is seen as the most flexible and successful item selection technique (Leite et al., 2008, Schultze & Eid, 2018). This success of the ACO results from the fact that it evaluates and optimizes the final model by assessing psychometric criteria in the scale rather than item-level information. The other algorithm, the tabu search (TS) algorithm (Drezner et al., 1999), looks at each short form item set created changing one item at a time. The main idea behind the TS procedure is to consistently identify the best short form currently selected by examining other short forms that are neighbors of the best short form available. If a neighboring short form under investigation fits better than the existing short form, it is selected as the most suitable new short form. If it does not fit with the existing short form, the neighboring short form examined is marked as "tabu" and these tabu items are placed in a list so that they are not reevaluated until certain criteria are met. The simulated annealing (SA) algorithm (Kirkpatrick et al., 1983), another algorithm aimed at developing a short form, "begins with a specific short form in which some items are randomly changed. The new short form is compared with the initial short form and the difference in model fit is calculated. At any time, if the new short form fits better than the existing short form, it is selected for use in the next iteration; otherwise, the new short form is chosen with probability equal to possibly a function of the difference in model fit and the difference between current and maximum iterations. This process continues until the maximum number of iterations is reached and the best short form final solution is found" (Raborn et al., 2020, p. 914). I chose the ACO, TS, and SA algorithms because these algorithms are the most well-structured metaheuristic algorithms in the literature (Marti et al., 2018).

Available Short Forms of the DERS

Kaufman et al. (2016) developed the short form of the DERS. With its 18-item and 6-factor structure, DERS-SF has the same factor structure as the long form of the original DERS. Moreira et al. (2020) tested DERS' short form (DERS-SF) and a modified version of DERS-SF's factor structure. For the adult sample, the correlated 6-factor model shows the best model fit of DERS-SF, but for the adolescent sample, the bifactor structure shows the best model fit. These findings indicate that both DERS and DERS-SF can fit 6-factor and bifactor models. Therefore, this study focused on this bifactor structure. Also, this study focused on Kaufman et al.'s (2016) version of DERS-SF, which has an 18-item and 6-factor structure. The superiority of this study compared to other DERS-SF development studies is that it uses metaheuristic algorithms. It was determined that all other studies (e.g., Kaufman et al., 2016; Moreira et al., 2020) used classical short form development techniques.

Findings of lower emotional expression among males versus females (Brody & Hall, 1993) suggest that males may have less emotional awareness than females. For this reason, I felt the need to include testing the measurement invariance according to gender as a hypothesis in the research. Also, studies suggest that self-compassion plays a crucial role in chronic pain experience with evidence indicating that self-compassion improves mental health through the enablement of adequate emotion regulation strategies (Diedrich et al., 2014). For this reason, the relationships between these two psychological constructs were also examined within the scope of concurrent validity.

In this context, the general purpose of this research is to create the short form of the Difficulties in Emotion Regulation Scale (DERS) by comparing the findings obtained from various metaheuristic algorithms. Within the framework of this general purpose, the correlated 6-factor and bifactor structures of the DERS will be tested with CFA, with items determined by metaheuristic algorithms. It was accepted that the algorithm with the highest model-data fit among the metaheuristic algorithms produced the most accurate result. The aim here is to benefit from the findings of the algorithm that produced the best finding, instead of comparing the findings of the algorithms. Then, measurement invariance by gender and DERS' correlation with the Self-Compassion Scale (SCS) was tested.

METHOD

Participants

The study group of the research consists of 398 students in various programs at Akdeniz University (Antalya, Turkey). Three hundred and sixteen of the participants are women and 82 of them are men. Although the average age is 22.16 (± 3.82), it ranges from 18 to 48. Being over the age of 18 and being a student of Akdeniz University were chosen as the inclusion criteria for the sample. For this purpose, participants were selected by criterion sampling technique. The measurement tool was communicated to the participants via online platforms. Informed consent was obtained from all individual participants included in the study. All procedures performed in the study 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.

Measures

Personal Information Form. The participants were asked for information that required a voluntary response, including gender, age, the state of having a sad emotional or physical life in childhood, parental attitude, mother's education level, and father's education level.

Difficulties in Emotion Regulation Scale (DERS). The adult version of the DERS was adapted into Turkish by Rugancı and Gençöz (2010). While examining the scale's factor structure for the adult sample, researchers excluded Item 10 from the Turkish version of the scale because of low factor loading and low correlation with the total scale. The exploratory factor analysis results also confirmed the 6-factor structure of the DERS. Additionally, the groups with high and low distress were successfully differentiated based on the DERS subscales (Rugancı & Gençöz, 2010). Also, Sarıtaş-Atalar et al. (2015) tested the 6-factor structure of the DERS among adolescents in a Turkish sample. In our study, we used Rugancı and Gençöz's (2010) translation of the scale. The Turkish short form of the DERS was adapted by Yiğit and Guzey Yiğit (2019). Bjureberg et al.'s (2016) DERS-SF was adapted in this study. Our study differs from this research because it aimed to develop a short form from the long form.

Self-Compassion Scale (SCS). The Self-Compassion Scale (SCS), which was developed by Neff (2003) and adapted into Turkish by Deniz et al. (2008), was used in the study. Originally SCS consists of 26 items and six subscales that include self-kindness (five items), self-judgment (five items), common humanity (four items), isolation (four items), mindfulness (four items), and over-identification (four items). Responses are given on a 5-point Likert scale ranging from 1 (*almost never*) to 5 (*almost always*). As a result of the adaptation of the SCS to the Turkish culture, a 24-item and unidimensional structure was determined, different from the original scale (Deniz et al., 2008). With this study, the SCS has a unidimensional model and a 24-item self-report measure.

Analysis

At this stage, the original DERS, which has 36 items and six correlated factors, and the possible short form alternatives of DERS-SF will be compared. These comparisons will take place with various metaheuristic algorithms.

Based on previous research (Marcoulides & Falk, 2018; Raborn et al., 2020), some tuning parameters are used for each metaheuristic algorithm. For ACO, 20 consecutive steps for convergence, 0.9 evaporation, 20 ants, and 1000 maximum steps for no improvement are tuned. For TS, 50 iterations and five tabu list size; for SA 1000 iterations are tuned for all factor structures. Tuning parameters affect the speed with which algorithms approach a solution, the criteria the algorithms use to determine the quality of the short form, and how many new short forms are allowed to be discovered (Raborn et al., 2020).

The model fit was checked using several fit indices, including the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the root-mean-square error of approximation (RMSEA). CFI and TLI values above .90 and .95 reflect acceptable and excellent fit, respectively, while RMSEA below or near .05 indicates an acceptable fit of data to a model (Hu & Bentler 1999). For all factor structures, omega coefficients as composite reliability are computed. Omega hierarchical (ωH) and omega hierarchical subscale (ωHS) are computed for bifactor structures. And omega hierarchical subscales (ωHSs) are computed for the 6-factor structure.

The scale's short form selection with ACO, TS, and SA was implemented with the ShortForm 0.4.6 package (Raborn & Leite, 2018). This package uses the lavaan package (Rosseel, 2012) to fit a unidimensional, 6-factor solution, and bifactor CFA analysis. Omega coefficients are computed from semTools package 0.5-4 (Jorgensen et al., 2022).

Within the scope of factorial validity, measurement invariance was tested according to gender. ShortForm 0.4.6 package (Raborn & Leite, 2018) was used for this purpose. Measurement invariance was evaluated through four nested levels: (i) configural invariance, (ii) metric invariance, (iii) scalar invariance, and (iv) strict invariance. Hu and Bentler's (1999) criteria were used for configural invariance. In the testing of nested models, it was decided according to the criteria of ΔCFI > –.01, ΔTLI > –.02, and $\Delta\chi^2$ ($p > .05$). $\Delta\chi^2$ is a test of the Satorra-Bentler χ^2 difference of the nested models, and ΔCFI and ΔTLI are the difference in CFI and TLI, respectively, between the two nested models.

The correlation of the scores obtained from different factor structures of the scale with the Self-Compassion Scale was determined. The Turkish version of the SCS has a unidimensional structure with 24 items. Several studies report a strong negative correlation between emotion regulation and self-compassion (Eichholz et al., 2020; Scoglio et al., 2018). Eichholz et al. (2020) reported a correlation of –.61 between emotion regulation and self-compassion and Scoglio et al. (2018) reported a correlation of –.71. All *R* codes are publicly available at https://osf.io/smyqk/

Results

The correlated 6-factor structure DERS long form was tested with CFA. Standardized regression coefficients of all 36 items vary between .32 and .87. Except for RMSEA, model-data fit shows poor model fit (TLI = .881; CFI = .891; RMSEA = .068). Omega coefficients are calculated as .91 for nonacceptance, .90 for the goals factor, .92 for the impulse factor, .92 for the strategies factor, .74 for the awareness factor, and .85 for the clarity factor. According to these findings, it can be said that the model-data fit of the correlated 6-factor structure is poor, but it has acceptable model-data fit and reliability values because it meets the lowest criteria according to RMSEA.

According to the findings obtained by applying CFA to the bifactor structure of the scale, standardized factor loadings to specific factors of all 36 items vary between .10 and .81; the general factor of all 36 items varies between .07 and .78. Model-data fit shows adequate model fit (TLI = .918; CFI = .929; RMSEA = .057). Omega coefficients are calculated as .25 for nonacceptance, .56 for the goals factor, .24 for the impulse factor, .22 for the strategies factor, .72 for the awareness factor, .81 for the clarity factor, and .75 for the general factor. According to these findings, it can be said that the model-data fit of the bifactor structure meets the minimum criteria and the model is confirmed. When the findings of the original form of the scale are examined, it can be said that the model-data fit of the bifactor model is higher and the factor structure of the DERS is more suitable for the bifactor model. Correlated 6-factor and bifactor structures of DERS were tested with CFA. Findings are given in Tables 1 and 2.

When the findings of the metaheuristic algorithms used in selecting the items to create the short form of the DERS were examined, it was determined that the 18 items and 6-factor structure created from the items selected by the ACO have the best psychometric properties. The factor structure with the highest average of standard factor loadings (.82), model-data fit, and reliability coefficients (.86) was obtained from the items determined by ACO. The factor structure obtained from the items selected by SA from other metacognitive algorithms also provided sufficient values in terms of all psychometric properties. However, the factor structure of the items determined by TS has moved away from perfect model-data fit, especially due to RMSEA and TLI values. It can be said that the 6-factor structure of DERS-SF has a perfect model-data fit (RMSEA = .042; CFI = .982; TLI = .976) according to the ACO algorithm and its reliability is quite high.

TABLE 1
Psychometric properties of DERS' correlated 6-factor structure

ACO		TS		SA	
Item	SFL	Item	SFL	Item	SFL
I29	**.86**	I29	**.83**	I29	**.84**
I23	**.80**	I11	**.67**	I23	**.81**
I25	**.82**	I23	**.81**	I21	**.74**
I26	**.85**	I13	**.77**	I13	**.78**
I33	**.83**	I26	**.84**	I26	**.83**
I18	**.86**	I33	**.86**	I33	**.83**
I27	**.85**	I3	**.65**	I32	**.82**
I14	**.86**	I32	**.83**	I14	**.88**
I19	**.89**	I27	**.84**	I19	**.87**
I35	**.78**	I16	**.81**	I31	**.66**
I16	**.82**	I31	**.68**	I16	**.86**
I31	**.67**	I36	**.83**	I15	**.83**
I2	**.79**	I6	**.74**	I6	**.73**
I6	**.77**	I8	**.84**	I34	**.53**
I8	**.81**	I10	**.35**	I10	**.39**
I7	**.83**	I5	**.57**	I4	**.85**
I4	**.85**	I4	**.75**	I1	**.74**
I1	**.75**	I1	**.80**	I7	**.83**
Model-data fit					
RMSEA	.042	RMSEA	.062	RMSEA	.047
CFI	.982	CFI	.950	CFI	.973
TLI	.976	TLI	.937	TLI	.966
Reliability (omega)					
Nonacceptance	.87	Nonacceptance	.82	Nonacceptance	.84
Goals	.89	Goals	.86	Goals	.86
Impulse	.90	Impulse	.84	Impulse	.89
Strategies	.80	Strategies	.82	Strategies	.84
Awareness	.83	Awareness	.70	Awareness	.56
Clarity	.85	Clarity	.77	Clarity	.85

Note. DERS = Difficulties in Emotion Regulation Scale; ACO = ant colony optimization; TS = tabu search; SA = simulated annealing. SFL = standardized factor loadings; factor loadings shown in bold are statistically significant at the .05 level. RMSEA = root-mean-square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index.

TABLE 2
Psychometric properties of DERS' bifactor structure

ACO			TS			SA		
Item	SFL	GSFL	Item	SFL	GSFL	Item	SFL	GSFL
I12	**.49**	**.52**	I25	**.38**	**.74**	I21	**.67**	**.56**
I29	**.40**	**.72**	I11	**.37**	**.62**	I23	**.23**	**.77**
I23	**.29**	**.77**	I23	**.17**	**.77**	I11	**.28**	**.61**
I18	**.67**	**.61**	I20	**.58**	**.43**	I20	**.58**	**.42**
I26	**.51**	**.67**	I13	**.58**	**.56**	I13	**.60**	**.54**
I33	**.41**	**.72**	I33	**.42**	**.67**	I33	**.46**	**.66**
I14	**.58**	**.63**	I14	**.70**	**.64**	I14	**.70**	**.63**
I19	**.62**	**.66**	I19	**.44**	**.72**	I19	**.48**	**.69**
I27	**.46**	**.71**	I3	**.20**	**.67**	I3	**.22**	**.66**
I16	**.57**	**.67**	I16	**.37**	**.73**	I36	**.44**	**.72**
I15	**.49**	**.66**	I31	**.25**	**.62**	I35	**.44**	**.70**
I31	**.31**	**.58**	I36	**.35**	**.75**	I15	**.35**	**.70**
I2	**.78**	**.17**	I6	**.76**	**.19**	I10	**.41**	.05
I6	**.75**	.12	I8	**.77**	**.18**	I6	**.67**	**.19**
I17	**.27**	.12	I34	**.47**	.12	I34	**.55**	.10
I4	**.69**	**.41**	I4	**.70**	**.36**	I7	**.77**	**.38**
I9	**.31**	**.66**	I9	**.36**	**.64**	I1	**.56**	**.47**
I1	**.64**	**.50**	I7	**.85**	**.35**	I9	**.35**	**.67**
Model-data fit								
RMSEA	.041		RMSEA	.044		RMSEA	.042	
CFI	.984		CFI	.980		CFI	.980	
TLI	.975		TLI	.970		TLI	.970	
Reliability (omega)								
Nonacceptance	.21		Nonacceptance	.13		Nonacceptance	.21	
Goals	.34		Goals	.38		Goals	.40	
Impulse	.37		Impulse	.27		Impulse	.30	
Strategies	.29		Strategies	.14		Strategies	.22	
Awareness	.63		Awareness	.68		Awareness	.55	
Clarity	.42		Clarity	.55		Clarity	.44	
DERS	.74		DERS	.77		DERS	.73	

Note. DERS = Difficulties in Emotion Regulation Scale; ACO = ant colony optimization; TS = tabu search; SA = simulated annealing. SFL = standardized factor loadings; GSFL = general factor's standardized factor loadings; factor loadings shown in bold are statistically significant at the .05 level. RMSEA = root-mean-square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index.

According to the findings obtained from the metaheuristic algorithms, it was determined that the average of the standard factor loadings of the specific factors (.51) and the standard factor loadings of the general factor (.55) obtained with the bifactor structure created with the items determined by the ACO algorithm were found to be the highest. In addition, the bifactor structure with the highest model-data fit and omega hierarchical subscale (HS) averages (.38) was obtained with the ACO algorithm. The algorithm with the highest omega hierarchy is TS. All algorithms succeeded in selecting the items that create the bifactor structure with excellent model-data fit. However, when the general psychometric properties are examined, it is possible to say that the ACO algorithm is the most successful method in creating short forms in the bifactor structure. It can be said that the bifactor structure of DERS-SF has a perfect model-data fit (RMSEA = .041; CFI = .984; TLI = .975) according to the ACO algorithm and its reliability is quite high.

According to Table 3, it can be said that most of the items (62.5%) in the DERS-SF developed with ACO are similar to Kaufman et al.'s (2016) DERS-SF items. Within the scope of factorial validity, the measurement invariance of the 6-factor structure, which has the highest model-data fit, was tested according to gender. Findings are included in Table 4.

All of the models met the criteria of ΔCFI > –.01, ΔTLI > –.02 and $\Delta\chi^2$(p > .05). All Satorra-Bentler chi-square differences are not significant at .05 level. The differences in all model-data fit were too small for all nested models. These results suggest that the 18-item 6-factor model of DERS meets the criteria for strict invariance.

TABLE 3
Item distributions of DERS-SF (ACO) and DERS-SF (Kaufman et al., 2016)

Factors	DERS-SF (ACO)	DERS-SF (Kaufman et al., 2016)
Strategies	I16	I35
	I15	I28
	I31	I16
Nonacceptance	I12	I12
	I29	I25
	I23	I29
Impulse	I14	I14
	I19	I21
	I27	I27
Goals	I18	I18
	I26	I26
	I33	I13
Awareness	I2	I8
	I6	I10
	I17	I2
Clarity	I4	I9
	I9	I5
	I1	I4

Note. DERS-SF = Difficulties in Emotion Regulation Scale-Short Form; ACO = ant colony optimization.

TABLE 4
Fit indices of measurement invariance models

Model	$\Delta\chi^2(df)$	RMSEA	CFI	ΔCFI	TLI	ΔTLI
Configural	–	.048	.977	–	.970	–
Metric	15.86 (12)	.047	.976	–.001	.970	.000
Scalar	10.60 (12)	.046	.976	.000	.972	.002
Strict	11.65 (18)	.039	.981	.005	.980	.008

Note. *df* = degrees of freedom; RMSEA = root-mean-square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index.

As the concurrent validity, the correlations between the 6-factor structure of the DERS and the Self-Compassion Scale, which has a unidimensional factor structure, were examined. The correlations between the total scores obtained from the Self-Compassion Scale and the nonacceptance, goals, impulse, strategies, awareness, and clarity factors of DERS are negative and vary between moderate to high (r = –.66, r = –.66, r = –.68, r = –.73, r = –.38, and r = –.50, respectively).

CONCLUSION

The model-data fit obtained from the correlated 6-factor structure (TLI = .976; CFI = .982; RMSEA = .042) and the bifactor structure (TLI = .975; CFI = .984; RMSEA = .041) created by the items selected by the ACO algorithms indicate a perfect fit. The values obtained are also very close to each other. Therefore, it was decided that the most appropriate short form should be decided by examining other findings. In this case, it was decided that the short form of the 6-factor structure would be the most appropriate since the standard factor loadings and reliability values of the correlated 6-factor structure formed by the items selected by the ACO algorithm were higher. Measurement invariance findings and correlation between other psychological traits are other important evidence. Other studies in the literature in which the psychometric properties of the 6-factor structure were determined to be the best were effective in this selection. This finding is similar to the studies of Kaufman et al. (2016) and Moreira et al. (2020). Both studies developed DERS-SF as a 6-factor structure. Only Moreira et al. (2020) found a higher model-data fit of the bifactor structure in the adolescent sample.

As a result of this study, DERS-SF was developed as six factors and 18 items (each factor has three items). Nonacceptance of emotional responses (nonacceptance) factor consists of Items 12, 23, and 29; difficulties engaging in goal-directed activities (goals) factor consists of Items 18, 26, and 33; impulse control difficulties (impulse) factor consists of Items 14, 19, and 27; limited access to emotion regulation strategies (strategies) factor consists of Items 15, 16, and 31; the lack of emotional awareness (awareness) factor consists of Items 2, 6, and 17; and the lack of emotional clarity (clarity) factor consists of Items 1, 4, and 9. Items 1, 2, 4, 6, 7, and 8 should be coded reversely. In the Turkish sample, with this short form, it will be possible to measure emotion regulation difficulties using less time and cost compared to the long form.

When compared to the fit values obtained from this study (TLI = .976; CFI = .982; RMSEA = .042) with model-data fit findings of Kaufman et al.'s (2016) study (CFI = .97; TLI = .96; RMSEA = .05) and Moreira et al.'s (2020) study (CFI = .954; TLI = .942; RMSEA = .046) from the adult sample, the success of the ACO algorithm in item selection emerges. In one of these two studies, Kaufman et al. (2016) followed a 3-step approach to short form development. Accordingly, in the first stage, factor loadings for each factor were listed in descending order. In the second step, weighting was made with factor loadings obtained from adult and adolescent samples. In the third stage, the "discrimination score" was obtained and the difference between loading an item on its specific factor and other factors was calculated. Moreira et al. (2020) used a short form structure developed by Kaufman et al. (2016). The fact that the short form has been developed with a higher model-data fit than other studies shows the power of metaheuristic algorithms.

With the selection of the 6-factor structure, it is necessary to examine the correlation between factors. Correlations between factors are moderate and high except for the awareness factor. The awareness factor shows a low correlation with other factors. These findings are similar to Moreira et al.'s (2020) and Osborne et al.'s (2017) studies.

DERS-SF shows full measurement invariance (strict invariance). Strict invariance is found, so the differences between gender groups' item responses are merely due to group differences. These findings are strong validity evidence. As further validity evidence, the correlations between the total score obtained from the Self-Compassion Scale and the nonacceptance, goals, impulse, strategies, awareness, and clarity factors of DERS-SF are negative and change from moderate to high. These findings are similar to Eichholz et al.'s (2020) and Scoglio et al.'s (2018) studies.

It is also necessary to compare the performance of ACO, TS, and SA algorithms, which are metaheuristic algorithms, in various factor structures. It was determined that the ACO algorithm, in a 6-factor structure, made more accurate item selection than the other algorithms. Perfect model-data fit has also been achieved with the SA algorithm. Only the TS algorithm has moved away from perfect model-data fit due to TLI and RMSEA values. In this case, SA and TS algorithms follow the ACO algorithm, respectively. In the bifactor structure, the structures created with the items determined by all algorithms show perfect fit. The model-data fits and omega coefficients obtained from each of the algorithms are very close to each other. Therefore, it can be said that there is no prominent algorithm in the bifactor structure. Finding the ACO algorithm as the best item selection algorithm is similar to the studies of Leite et al. (2008) and Schultze and Eid (2018). TS and SA algorithms show similar findings.

This study has some limitations. In this study, the genetic algorithm, one of the metaheuristic algorithms, is not used. I would have liked to use genetic algorithm also in this study, but possible short form findings could not be produced. The possible reason for this is the insufficient sample size. For this reason, it is recommended to carry out studies in which this algorithm is also applied by using a sufficient sample size. Simulation-based studies in which these algorithms are compared with each other are required. Few studies with this aim exist in the literature. For the practitioners, it is recommended to use the short form developed in this study instead of the long form of the DERS. DERS-SF has half fewer items and is a much superior measurement tool in terms of psychometric properties.

References

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review*, *30*, 217-237. https://doi.org/10.1016/j.cpr.2009.11.004

Bardeen, J. R., Fergus, T. A., & Orcutt, H. K. (2012). An examination of the latent structure of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*, *34*, 382-392. https://psycnet.apa.org/doi/10.1007/s10862-012-9280-y

Benfer, N., Bardeen, J. R., Fergus, T. A., & Rogers, T. A. (2019). Factor structure and incremental validity of the original and modified versions of the difficulties in emotion regulation scale. *Journal of Personality Assessment*, *101*(6), 598-608. http://dx.doi.org/10.1080/00223891.2018.1492927

Bjureberg, J., Ljotsson, B., Tull, M. T., Hedman, E., Sahlin, H., Lundh, L. G., Bjärehed, J., DiLillo, D., Messman-More, T., Gumpert, C. H., & 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

Brody, L. R., & Hall, J. A. (1993). Gender and emotion. In J. Haviland & M. Lewis (Eds.), *Handbook of emotions* (pp. 447-460). Guilford Press.

Byrne, G. J., & Pachana, N. A. (2011). Development and validation of a short form of the Geriatric Anxiety Inventory-the GAI-SF. *International Psychogeriatrics*, *23*(1), 125-131. https://doi.org/10.1017/S1041610210001237

Catanzaro, S. J., & Mearns, J. (1990). Measuring generalized expectancies for negative mood regulation: Initial scale development and implications. *Journal of Personality Assessment*, *54,* 546-563. https://doi.org/10.1080/00223891.1990.9674019

Colorni, A., Dorigo, M., & Maniezzo, V. (1991). Distributed optimization by ant colonies. In F. Varela & P. Bourgine (Eds.), *Proceedings of the European Conference on Artificial Life, ECAL '91* (pp. 134-142). Elsevier Publishing.

Dan-Glauser, E. S., & Scherer, K. R. (2013). The Difficulties in Emotion Regulation Scale (DERS): Factor structure and consistency of a French translation. *Swiss Journal of Psychology*, *72*(1), 5-11. https://doi.org/10.1024/1421-0185/a000093

Deniz, M., Kesici, Ş., & Sümer, A. S. (2008). The validity and reliability of the Turkish version of the Self-Compassion Scale. *Social Behavior and Personality: An International Journal*, *36*(9), 1151-1160. https://doi.org/10.2224/sbp.2008.36.9.1151

Diedrich, A., Grant, M., Hofmann, S. G., Hiller, W., & Berking, M. (2014). Self-compassion as an emotion regulation strategy in major depressive disorder. *Behaviour Research and Therapy*, *58*, 43-51. https://doi.org/10.1016/j.jsxm.2019.11.266

Dorigo, M., Birattari, M., & Stutzle, T. (2006). Ant colony optimization. *IEEE Computational Intelligence Magazine*, *1*(4), 28-39. https://doi.org/10.1109/MCI.2006.329691

Drezner, Z., Marcoulides, G. A., & Salhi, S. (1999). Tabu search model selection in multiple regression analysis. *Communications of Statistics*, *28*(2), 349-367. https://doi.org/10.1080/03610919908813553

Ehring, T., & Quack, D. (2010). Emotion regulation difficulties in trauma survivors: The role of trauma type and PTSD symptom severity. *Behavior Therapy*, *41*, 587-598. http://dx.doi.org/10.1016/j.beth.2010.04.004

Eichholz, A., Schwartz, C., Meule, A., Heese, J., Neumüller, J., & Voderholzer, U. (2020). Self-compassion and emotion regulation difficulties in obsessive-compulsive disorder. *Clinical Psychology & Psychotherapy*, *27*(5), 630-639. https://doi.org/10.1002/cpp.2451

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*(1), 41-54. https://doi.org/10.1023/B:JOBA.0000007455.08539.94

Gratz, K. L., Rosenthal, M., Tull, M. T., Lejuez, C. W., & Gunderson, J. G. (2006). An experimental investigation of emotion dysregulation in borderline personality disorder. *Personality Disorders: Theory, Research, and Treatment*, *1*, 18-26. https://doi.org/10.1037/0021-843x.115.4.850

Gross, J. J., & Munoz, R. F. (1995). Emotion regulation and mental health. *Clinical Psychology: Science and Practice*, *2*, 151-164. https://psycnet.apa.org/doi/10.1111/j.1468-2850.1995.tb00036.x

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., Rosseel, Y., Miller, P., Quick, C., Garnier-Villarreal, M., Selig, J., Boulton, A., Preacher, K., Coffman, D., Rhemtulla, M., Robitzsch, A., Enders, C., Arslan, R., Clinton, B., Panko, P., Merkle, E., Chesnut, S., … Johnson, A. R. (2022). *semTools: Useful tools for structural equation modeling. R package* (Version 0.5-4). Retrieved from https://CRAN.R-project.org/package=semTools

Kaufman, E. A., Xia, M., Fosco, G., Yaptangco, M., Skidmore, C. R., & Crowell, S. E. (2016). The Difficulties in Emotion Regulation Scale-Short Form (DERS-SF): Validation and replication in adolescent and adult samples. *Journal of Psychopathology and Behavioral Assessment*, *38*, 443-455. https://doi.org/10.1007/s10862-015-9529-3

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science, 220*(4598), 671-680. https://doi.org/10.1126/science.220.4598.671

Kleka, P., & Soroko, E. (2018). How to avoid the sins of questionnaires abridgement? *Survey Research Methods*, *12*(2), 147-160. http://dx.doi.org/10.18148/srm/2018.v12i2.7224

Kopp, C. B. (1989). Regulation of distress and negative emotions: A developmental view. *Developmental Psychology*, *25*, 343-354. https://doi.org/10.1037/0012-1649.25.3.343

Kruyen, P. M., Emons, W. H., & Sijtsma, K. (2013). On the shortcomings of shortened tests: A literature review. *International Journal of Testing*, *13*(3), 223-248. https://doi.org/10.1080/15305058.2012.703734

Leite, W. L., Huang, I. C., & Marcoulides, G. A. (2008). Item selection for the development of short forms of scales using an ant colony optimization algorithm. *Multivariate Behavioral Research*, *43*, 411-431. https://doi.org/10.1080/00273170802285743
Levy, P. (1968). Short-form tests: A methodological review. *Psychological Bulletin*, *69*(6), 410.
Marcoulides, K. M., & Falk, C. (2018). Model specification searches in structural equation modeling with R. *Structural Equation Modeling*, *25*(3), 484-491. https://doi.org/10.1080/10705511.2017.1409074
Marti, R., Pardalos, P. M., & Resende, M. G. C. (2018). *Handbook of heuristics*. Springer.
Moreira, H., Gouveia, M. J., & Canavarro, M. C. (2020). A bifactor analysis of the Difficulties in Emotion Regulation Scale-Short Form (DERS-SF) in a sample of adolescents and adults. *Current Psychology*, *41*, 757-782. https://doi.org/10.1007/s12144-019-00602-5
Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity*, *2*(3), 223-250. https://doi.org/10.1080/15298860390209035
Neumann, A., van Lier, P. A., Gratz, K. L., & Koot, H. M. (2010). Multidimensional assessment of emotion regulation difficulties in adolescents using the difficulties in emotion regulation scale. *Assessment*, *17*(1), 138-149. https://doi.org/10.1177/1073191109349579
Noble, W., Jensen, N. S., Naylor, G., Bhullar, N., & Akeroyd, M. A. (2013). A short form of the Speech, Spatial and Qualities of Hearing scale suitable for clinical use: The SSQ12. *International Journal of Audiology*, *52*(6), 409-412. https://doi.org/10.3109/14992027.2013.781278
Osborne, T. L., Michonski, J., Sayrs, 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 nonsuicidal self-injury in adolescent inpatients. *Journal of Psychopathology and Behavioral Assessment*, *34*(3), 393-404. https://psycnet.apa.org/doi/10.1007/s10862-012-9292-7
Raborn, A. W., & Leite, W. L. (2018). ShortForm: An R package to select scale short forms with the ant colony optimization algorithm. *Applied Psychological Measurement*, *42*(6), 516-517. https://doi.org/10.1177%2F0146621617752993
Raborn, A. W., Leite, W. L., & Marcoulides, K. M. (2020). A comparison of metaheuristic optimization algorithms for scale short-form development. *Educational and Psychological Measurement*, *80*(5), 910-931. https://doi.org/10.1177%2F0013164420906600
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
Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02
Ruganci, R. N., & Gençöz, T. (2010). Psychometric properties of a Turkish version of the Difficulties in Emotion Regulation Scale. *Journal of Clinical Psychology*, *66*(4), 442-455. https://doi.org/10.1002/jclp.20665
Salovey, P., Mayer, J. D., Goldman, S. L., Turvey, C., & Palfai, T. P. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the trait meta-mood scale. In J. W. Pennebaker (Ed.), *Emotion, disclosure, and health* (pp.125-154). American Psychological Association. https://doi.org/10.1037/10182-006
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*(1), 12-19. https://psycnet.apa.org/doi/10.1027/1015-5759/a000199
Schroeders, U., Wilhelm, O., & Olaru, G. (2016). Meta-heuristics in short scale construction: Ant colony optimization and genetic algorithm. *PLoS One*, *11*(11). Article e0167110. https://doi.org/10.1371/journal.pone.0167110
Scoglio, A. A., Rudat, D. A., Garvert, D., Jarmolowski, M., Jackson, C., & Herman, J. L. (2018). Self-compassion and responses to trauma: The role of emotion regulation. *Journal of Interpersonal Violence*, *33*(13), 2016-2036. https://doi.org/10.1177/0886260515622296
Schultze, M., & Eid, M. (2018). Identifying measurement invariant item sets in cross-cultural settings using an automated item selection procedure. *Methodology*, *14*(4), 176-187. https://doi.org/10.1027/1614-2241/a000155
Tull, M. T., Barrett, H. M., McMillan, E. S. & Roemer, L. (2007). A preliminary investigation of the relationship between emotion regulation difficulties and posttraumatic stress symptoms. *Behavior Therapy*, *38*(3), 303-313. https://doi.org/10.1016/j.beth.2006.10.001
Yiğit, İ., & Guzey Yiğit, M. (2019). Psychometric properties of Turkish version of Difficulties in Emotion Regulation Scale-Brief Form (DERS-16). *Current Psychology*, *38*(6), 1503-1511. https://link.springer.com/article/10.1007/s12144-017-9712-7