

Nordic Journal of Psychiatry



ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ipsc20

Validation of the Swedish version of the Reynolds Adolescent Depression Scale second edition (RADS-2) in a normative sample

Ida Blomqvist, Erik Ekbäck, Inga Dennhag & Eva Henje

To cite this article: Ida Blomqvist, Erik Ekbäck, Inga Dennhag & Eva Henje (2021) Validation of the Swedish version of the Reynolds Adolescent Depression Scale second edition (RADS-2) in a normative sample, Nordic Journal of Psychiatry, 75:4, 292-300, DOI: 10.1080/08039488.2020.1850858

To link to this article: https://doi.org/10.1080/08039488.2020.1850858

9	© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
	Published online: 28 Nov 2020.
	Submit your article to this journal 🗗
hh	Article views: 426
Q	View related articles ☑
CrossMark	View Crossmark data ௴



ARTICLE



Validation of the Swedish version of the Reynolds Adolescent Depression Scale second edition (RADS-2) in a normative sample

Ida Blomqvist (D), Erik Ekbäck (D), Inga Dennhag (D) and Eva Henje (D)

Department of Clinical Science, Child and Adolescent Psychiatry, Umea University, Umea, Sweden

ABSTRACT

Background: Due to the sharp global increase in prevalence of adolescent major depressive disorder as defined by the Diagnostic and Statistical Manual of Mental Disorders, we need internationally validated tools for multi-dimensional assessment. Reynolds Adolescent Depression Scale second edition (RADS-2) measures dysphoric mood, anhedonia/negative affect, negative self-evaluation and somatic complaints and is widely used internationally, but not yet available in Swedish.

Aim: The aim of this study is to test the psychometric characteristics of the Swedish version of RADS-2 in a normative sample.

Material and method: Data was gathered from junior and high school students in Northern Sweden (N = 637). We performed: 1. Confirmatory factor analysis to examine the 4-factor structure proposed by Reynolds, 2. Measurement invariance analysis for sex (girls, boys) and age group (12-15 years, 16-20 years). 3. Reliability testing and 4. Tests for concurrent, discriminant and convergent validity using Beck's Youth Inventories of Emotional and Social Impairment Depression and Anger subscales, the Patient Reported Outcome Measurements Information System, Anxiety and Friends subscales and the World Health Organization Wellness Index.

Results: The sample consisted of n = 637 students (n = 389 girls and n = 248 boys), mean age 15.73 (SD = 1.76); 12-20 years. The 4-factor structure was confirmed, as well as measurement invariance for sex and age group. Reliability was acceptable to excellent for all subscales and RADS-2 total scale. Concurrent, convergent and discriminant validity was good.

Conclusion: The Swedish version of RADS-2 showed acceptable reliability and validity in a Swedish normative sample.

ARTICLE HISTORY

Received 15 April 2020 Revised 14 October 2020 Accepted 9 November 2020

KEYWORDS

Adolescents; reliability; validity; confirmatory factor analysis; measurement invariance

Introduction

At present, major depressive disorder (MDD) accounts for a major part of the global disease burden in adolescents and young adults [1]. The WHO predicts that it will be the leading cause of the burden of disease worldwide by 2030 [2]. Early onset MDD increases the risk of recurrent episodes, a more serious course of the disease and suicide [3-5]. It is associated with increased morbidity in somatic diseases, such as cardiovascular conditions [6] and cancer [7], with increased mortality independent of suicide [8]. The Diagnostic and Statistical Manual of Mental Disorders (DSM) developed by the American Psychiatric Association is normally used to diagnose MDD in adolescents. Even though the DSM has many advantages such as the weighting of functioning level and severity assessment for mental disorders, it has been critiqued for a low diagnostic validity and unclear diagnostic boundaries especially for adolescent MDD [9-13].

The Research Domain Criteria (RDoC), an initiative of the National Institute of Health (NIMH), were created to aid the development of individualized precision medicine for mental health and 'help identify new targets for treatment development, detect subgroups for treatment selection, and provide a better match between research findings and clinical decision making' [14]. The structure of the RDoC is described as 'a matrix in which the rows represent various constructs grouped hierarchically into broad domains of function'. 'The columns of the matrix denote different levels of analysis, from genetic, molecular, and cellular levels, proceeding to the circuit-level', 'and on to the level of the individual, family environment, and social context' [14].

The Reynolds Adolescent Depression Scale second edition (RADS-2) is a well validated self-report questionnaire that measures four dimensions of adolescent MDD: dysphoric mood, anhedonia/negative affect, negative self-evaluation and somatic complaints, while it is still compatible with the DSM and the International Classification of Disease (ICD) systems [15]. Therefore, it constitutes a useful tool to identify dimensions of psychopathology according to the constructs and domains of function of the RDoC matrix in adolescent depression research. It is used both as a clinical measure to





aid diagnostics and to measure treatment effect, as well as in clinical research on adolescent MDD in several countries [16–21]. The RADS-2 has demonstrated strong psychometric characteristics and good to excellent concurrent, convergent and discriminative validity in international studies with clinical [22] and non-clinical samples of adolescents [16,18]. Confirmatory factor analyses have supported the 4-factor structure of RADS-2 [16,18,22].

The reliability of the RADS-2 and its subscales is also supported in the literature. In a large sample (N = 9052) a high Cronbach's alpha was found (0.93) for the RADS-2 total score, with subscales ranging from 0.80 to 0.87 [15]. A potential weakness of the RADS-2 questionnaire is that it does not specifically address the RDoC construct of attention. Attention deficit is a symptom criteria of MDD, that has been found to predict high scores in other depression inventories in this age-group, such as the Mood and Feelings Questionnaire (MFQ) [23]. On the other hand, attention deficit is a symptom criterion not specific to MDD, but also part of other psychiatric diagnoses, for example generalized anxiety disorder (GAD), attention deficit hyperactivity disorder (ADHD) and post-traumatic stress disorder (PTSD), which is a plausible reason for its omission from the RADS-2.

RADS-2 has not yet been translated and validated in Swedish and the aim of this study is to test the psychometric properties of the Swedish version of RADS-2 in a normative sample, this as a step towards a future clinical validation following the validation processes of the RADS-2 in other countries [16,18].

Materials and method

Participants

Participants were recruited to this convenience sample from four junior and high schools from different socioeconomic backgrounds in both suburb and rural areas in northern Sweden and included students from different school-programs such as natural science, social science, media and the arts. Inclusion criteria were 1. Being a student in one of the chosen schools/classes and 2. Being between 12 and 20 years of age. Exclusion criteria were 1. Non-fluency in written Swedish and 2. Inability to complete online forms (e.g. absence on the day of data collection or severe dyslexia). 897 students were asked to participate and n = 637 (71%) of them agreed to participate in the study, mean age 15.73 (SD

A subset of the total sample (53%, n = 338), consented to and completed the same questionnaires three weeks later at home, to obtain data on test-retest reliability. The mean age was 15.4 years (SD = 1.68). A presentation of the age- and sex-distributions of the two samples is given in Table 1.

In the total sample, 70% of the participants were living with both parents and most participants (88%) were born in Sweden. To estimate the socioeconomic status of the participants' households a Swedish socioeconomic classification system was used [24,25]. The distribution of the socioeconomic classification of the participants parents was as follows; 17.20% workers, 28.40% assistant and intermediate

Table 1. Age and sex distributions for the full sample and retest sample.

	Full sample, n (%)	Retest sample, n (%)
Age (years)		
12	12 (1.9)	9 (2.7)
13	74 (11.6)	51 (15.1)
14	73 (11.5)	36 (10.7)
15	136 (21.4)	87 (25.7)
16	109 (17.1)	53 (15.7)
17	117 (18.4)	60 (17.8)
18	91 (14.3)	41 (12.1)
19	20 (3.1)	1 (0.2)
20	5 (0.7)	0
total	637 (100)	338 (100)
Sex		
Male	248 (39)	108 (32)
Female	389 (61)	230 (68)
total	637 (100)	338 (100)

non-manual workers, 32.10% professionals, civil servants, and executives, 7.10% self-employed of various kinds, and 15.20% unknown.

Procedure

Permission was granted from the publishing company to translate RADS-2 into Swedish. In order to guarantee semantic and content equivalence and in accordance with recommended procedure for translating research instruments, a back-translation method was used [26]. Pilot testing of the questionnaire was conducted to test the overall features of the questionnaire as well as expected time consumption. The instrument was finally approved by the publisher. The study was approved by the Regional Ethics Board in Umeå.

The Principal Investigator (PI) of the study contacted the education-director of the municipality, who referred the PI to eight school-principals in the region. All of them were contacted in person and four agreed to participate in the study. The PI was then allowed to contact the classroom teachers of these schools and instructed them to inform the students about the study. Data was collected during school hours. Research assistants informed the students about the study in detail and gathered individual written consent for participation. Parental consent was gathered for students under the age of 15. All data-collection occasions were scheduled for one hour and supervised by a research assistant or the Pl. The majority of students completed the questionnaires in about 30-45 min, additional measures than the ones presented here were also administered [27]. The students could ask questions during the completion of the forms if necessary and a light snack was offered in the middle. The completion order of the questionnaires was altered between classrooms and between testing occasions. Data was collected during 2018 and 2019.

Instruments

Reynolds Adolescent Depression Scale second edition (RADS-2) [15] consists of 30 brief self-statements on a 4-point scale ranging from 'Almost never' to 'Most of the time'. The items are divided into four subscales/dimensions (dysphoric mood, anhedonia/negative affect, negative self-evaluation and



somatic complaints). The anhedonia/negative affect items are formulated with positive questions such as 'I feel happy' and thus are reversely coded. Higher scores indicate higher symptom severity and the scale has a theoretical raw score between 30 and 120 [15].

Instruments used for validation

The Beck Youth Inventories of Emotional and Social Impairment [28] - specifically the subscales of Depression (BYI-D) and Anger (BYI-A), each consisting of 20 brief selfstatement-questions on a 4-point scale ranging from 'Never' to 'Always' and each with a theoretical raw score ranging from 0 to 60 [28]. Higher scores indicate higher symptom severity. The BYI-D scale is extensively used in Swedish Child and Adolescent Depression clinics [29] and recommended by the Swedish Agency for Health Technology Assessment and Assessment of social services [30] as a scale to use when screening for MDD among adolescents. Cronbach's alpha in the current sample was 0.94 (95% CI [0.93, 0.95]) for BYI-D and 0.93 (95% CI [0.92, 0.94]) for BYI-A.

The World Health Organization Wellness Index (WHO-5) consists of 5 salutogenic self-statements. Question 1 for example is worded: 'Over the past two weeks I have felt cheerful and in good spirits'. Questions are answered on a 6point scale ranging from 'All of the time' to 'At no time', with a theoretical raw score ranging from 0 to 25. Higher scores indicate better well-being and a total score below 13 indicates poor well-being. The scale has adequate validity both as a screening tool for depression and as an outcome measure in clinical trials and has been applied successfully across a wide range of study fields [31]. It has also been validated in depressed adolescents in Sweden [32]. Cronbach's alpha in the current sample was 0.87 (95% CI [0.86, 0.89]).

The Patient Reported Outcome Measurements Information System (PROMIS) consists of item banks for various health and life-style dimensions that were developed to advance the science and application of patient reported outcomes. In this study the item banks for Anxiety (PROMIS Anxiety) and Peer relationships (PROMIS Friend) were used. The guestions are worded in past tense starting 'in the last 7 days ...', and PROMIS Anxiety consist of 15 questions (e.g. 'I felt like something awful might happen') and PROMIS Friend consists of 10 questions (e.g. 'I was able to count on my friends'), both on a 5-point scale, with possible answers ranging from 'Never' to 'Almost always' [33]. Theoretical score ranges are 0-60 for PROMIS anxiety and 0-40 for PROMIS friend and higher scores indicate higher anxiety-levels and better peer relationships respectively. Cronbach's alpha in the current sample was 0.92 (95% CI [0.91, 0.93]) for PROMIS Anxiety and 0.92 (95% CI [0.91, 0.93]) for PROMIS Friend.

Data analysis

Data was first analysed with descriptive statistics. The means and standard deviations for the items, subscales and total scores are reported separately for each sex (girls, boys) and age-group (12-15 years, 16-20 years), as well as for the whole

sample. Since the data was ordinal data Mann-Whitney U test was used to test mean differences between groups.

A confirmatory factor analysis (CFA) with a four-factor correlated model was performed to test the model proposed by Reynolds and previously confirmed in other versions of the scale [18]. The following indices were used to test the goodness-of-fit of the model; the Chi Square (χ^2) , the comparative fit index (CFI), The Tucker-Lewis index (TLI) and the root mean square error approximation (RMSEA). Acceptable fit for the model was determined by RMSEA < 0.06 absolute below 0.08, TLI and CFI >0.95 for excellent fit and >0.90 for acceptable fit [34-36]. Since the distribution was non-normal and with ordinal scale variables on a 4-point scale the default for ordered variables in R, the robust diagonal weighted least square (DWLSSS) estimator was used [37,38].

Measurement invariance according to Svetina Rutkowski [39] was tested in order to establish that the scale measures the same for both the older and younger age group and for both sexes. Since the observed means equal the intercept/thresholds of the variable added together with the factor loadings times the factor score, it is in theory possible for the intercept/thresholds to be unequal resulting in elevated or attenuated observed means for different groups giving biased observed means [40]. When measurement invariance holds it allows for the interpretation that the mean differences between groups are due to actual differences (and not because the questions in the scale are conceived differently in the different groups). Measurement invariance (MI) was tested with a forward procedure, first establishing a configural model for the different groups and thereafter step by step constraining thresholds and factor loadings to find evidence for metric and scalar measurement invariance. The model fit for the metric model was compared with model fit for the configural model, and the model fit for the scalar model was compared with the model fit for the metric model. Since the Chi square is sensitive to sample size, Δ CFI and Δ RSMEA were evaluated, and also Satorra-Bentler test was performed. If scalar measurement invariance holds it is possible to evaluate latent mean differences. To establish measurement invariance the following goodness of fit indices were used; $\Delta RSMEA \leq 0.05$ and significant $\Delta \chi^2$ and CFI \geq -0.004 for metric invariance and; $\Delta RSMEA \leq$ 0.01 and significant $\Delta \chi^2$ and CFI ≥ -0.002 [39,41,42], as well as a negative Satorra-Bentler test indicating that the null hypothesis fails and that there is no difference between the models.

Reliability was tested with Cronbach's alpha, which is extensively used as a measure of internal consistency even though during recent years its usefulness has been called into question [43], however its credibility in the field is still considered to be strong [44]. Also the comparative analyses of different measures of reliability showed the risk of underestimated values of Cronbach alpha when assumptions are violated [43]. Cronbach's alpha >0.7 was classified as 'acceptable', >0.8 as 'good', and >0.9 as 'excellent' [45].

For test-retest statistics Pearson's correlation coefficient (Pearsons r) was used. Test-retest correlations were interpreted as follows: 0.60-0.69 is questionable, 0.70-0.79 is acceptable, and 0.80-0.89 is good reliability [46,47].



Concurrent validity is shown if a measure of the same construct, in this case depression, has a high correlation with the scales currently validated and was tested by correlation estimates between RADS-2 and BYI-D. Convergent validity is considered to be established when a measure of a similar construct is correlated and was tested with correlation estimates between RADS-2 and PROMIS Anxiety, BYI-A subscale and WHO-5. Finally, discriminant validity is to test whether a measure of a different construct discriminates to the construct of the scale validated. In this study discriminant validity was considered to be established if the correlation between RADS-2 and PROMIS Friend was moderate. Concurrent, discriminant and convergent validity were examined with Pearson's r and 0.1-0.29 is considered to be a small correlation, 0.3-0.49 a medium correlation and 0.50 and above is considered to be a large correlation [48].

Data was analysed with SPSS 26 and R [49], using the Lavaan package for structural equation modelling version 0.6-3 [50].

Results

Descriptive statistics

Missing values ranged from 0 to 0.8% for all items of RADS-2, missing values analysis with Little's MCAR test for RADS-2 variables was non-significant (p < 0.497, Chi-Square =804.63, df =802), multiple imputations with FCS method in SPSS was used.

For item and subscale means and standard deviations of RADS-2 for the whole sample and reported separately for sex (girls, boys) and age groups (12-15 years, 16-20 years), see Table 2. Mean score for the whole sample was 59.51 (SD =15.78). Girls and 16- to 20-year-olds (both boys and girls) had significantly higher scores for Dysphoric mood, Negative selfevaluation, Somatic complaints subscale and the RADS-2 total scale. Corrected item-total correlation statistics for the items and their respective subscales are also reported, with all items being above 0.3.

Factor structure

Standardized factor loadings for the items separated on sex (girls, boys) and age groups (12-15 years, 16-20 years) and the total sample are shown in Table 3. The factor loadings for the total sample had a range from 0.41 (item 23, reduced speech on the anhedonia/negative affect subscale) to 0.95 (item 20, self-deprecation on the negative self-evaluation subscale) (see Table 3).

The model fit for the correlated 4-factor model for the total sample, chi square (sensitive to sample size) was significant, χ^2 (399) = 1738.61, p < 0.001, but the other fit indices were found to be acceptable CFI = 0.945, TLI = 0.940, RMSEA = 0.072 (90% CI [0.069-0.076]), the individual confirmatory factor analyses for sex (girls, boys) and age group (12–15 years, 16–20 years) yielded similar fit indices (see Table 4).

Measurement invariance (MI)

Measurement invariance was further tested and configural, metric and scalar measurement invariance was shown, see Table 4. RADS-2 was invariant across groups examined i.e. sex (girls, boys) and age group (12-15 years, 16.20 years) for all the above described levels of MI (see Table 5).

Reliability

All reliability measures were within acceptable to excellent range for all subscales and the RADS-2 total scale with Cronbach's alpha for dysphoric mood 0.87 (95% CI [0.85, 0.88]); anhedonia/negative affect 0.77 (0.74, 0.79); negative self-evaluation 0.87 (95% CI [0.86, 0.89]); somatic complaints 0.80 (95% CI [0.78, 0.82]); and RADS-2 total scale 0.93 (95% CI [0.93, 0.94). Test-retest with Pearsons r indicated questionable to good correlations, for dysphoric mood (0.82); anhedonia/negative affect (0.67); negative self-evaluation (0.80); somatic complaints (0.79); and RADS-2 total scale (0.86), all at p < 0.001 (2-tailed).

Convergent, concurrent and discriminant validity

Correlations between subscales and RADS-2 total scale as well as the scales used for cross-validation are shown in Table 6. The lowest correlations were found between dysphoric mood and anhedonia/negative affect (0.36, p < 0.01) and the highest correlation was found between negative self-evaluation and RADS-2 total scale (0.92, p < 0.01).

Correlations between RADS-2 and BYI-D ranged from 0.56 (anhedonia/negative affect) to 0.88 (RADS-2 total scale) thus indicating concurrent validity. Convergent validity was showed with correlations with PROMIS Anxiety ranging from 0.43 (anhedonia/negative affect) to 0.70 (RADS-2 total scale), BYI-A correlations ranging from 0.46 (anhedonia/negative affect) to 0.73 (RADS-2 total scale) and WHO-5 correlations ranging from 0.48 (anhedonia/negative affect) to 0.72 (RADS-2 total scale). Discriminant validity was showed with PROMIS FRIEND; with correlations ranging from -0.38 (somatic complaints) to -0.50 (anhedonia/negative affect) (for details see Table 6).

Discussion

The aim of this study was to test the psychometric properties of the Swedish version of the RADS-2 in a normative sample. RADS-2 is an internationally established multi-dimensional measure of adolescent depression that is compatible with the DSM- and ICD systems and not yet translated and used in Sweden.

The main finding of the study was that the RADS-2 showed evidence of being reliable in conformity with previous studies of non-clinical samples i.e. Reynolds with colleagues [15], since all Cronbach alpha values ranged from acceptable (anhedonia/negative affect subscale) to excellent (RADS-2 total scale). This makes it a reliable instrument to use in a Swedish community setting. Evidence was found for

Table 2. Item means, standard deviations for the Swedish RADS-2 subscales and the total scale for girls, boys, age groups, and total sample.

		Girls (N = 389)	= 389)	Bovs (N=248)	= 248)		12-15 vears (N=295)	(N = 295)	16–20 vears (N = 342)	(N = 342)		Total sample (N = 637)	(N = 637)	
Subscales	Abbreviation	_ ≥	S	\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	SD	z-score	_ ≥	SD	_ ≥	S	z-score	. ∑	S	Item – Subscale statistics
Dysphoric mood Items		19 17	4 95	15.25	5 14	-9.01**	16.42	5.56	1871	4 97	-551**	17.65	5 37	
2.	School anxiety	3.01	0.85	2.36	0.99		2.61	1.03	2.89	0.87	!	2.76	96.0	09:0
ű.	Loneliness	2.35	0.94	2.04	1.00		2.04	0.97	2.39	0.94		2.23	0.97	0.59
9.	Social withdrawal	5.09	1.01	1.79	0.91		1.87	0.99	2.06	0.98		1.97	0.99	0.50
7.	Sadness	2.62	0.81	1.99	0.87		2.22	0.95	2.51	0.81		2.38	0.88	0.78
8.	Crying	2.36	0.87	1.55	0.75		1.91	0.93	2.16	0.88		2.05	0.91	89.0
16.	Irritability	2.34	98.0	1.97	0.87		2.12	0.91	2.26	0.85		2.19	0.88	0.59
21.	Self-pity	1.91	0.87	1.58	98.0		1.58	0.81	1.96	06.0		1.78	0.88	0.52
26.	Worry	2.49	0.92	1.98	1.02		2.07	1.03	2.48	0.92		2.29	0.99	0.70
Anhedonia/ negative affect Items		10.98	3.59	10.82	3.42	$-0.36\mathrm{ns}$	10.81	3.47	11.01	3.57	-0.69 ns	10.92	3.52	
	Reduced affect	1.49	69.0	1.42	0.67		1.39	99.0	1.51	0.70		1.46	89.0	0.59
5.	Self-worth	1.93	0.87	1.80	98.0		1.85	0.88	1.90	98.0		1.88	0.87	0.51
10.	Anhedonia-peers	1.31	0.64	1.33	0.67		1.29	0.63	1.34	0.67		1.32	99.0	0.49
12.	Low self-worth	1.59	0.78	1.74	0.93		1.65	98.0	1.65	0.84		1.65	0.84	0.51
23.	Reduced speech	1.80	0.91	1.81	0.94		1.84	0.94	1.78	0.91		1.81	0.92	0.46
25.	Anhedonia-general	1.37	0.68	1.37	99.0		1.35	0.64	1.38	0.70		1.37	0.67	0.50
29.	Appetite disturbance	1.50	0.81	1.35	0.70		1.44	0.77	1.44	0.79		1.44	0.78	0.39
Negative self-evaluation Items		15.57	5.49	12.93	4.83	-6.31**	14.04	5.43	14.98	5.33	-2.56*	14.54	5.39	
4	Feelings of rejection	1.41	0.83	1.30	0.74		1.31	0.73	1.43	0.84		1.37	0.80	0.34
.6	Worthlessness	1.95	0.95	1.59	0.87		1.75	0.91	1.87	96.0		1.81	0.94	69.0
13.	Discouragement	1.89	1.01	1.53	0.85		1.72	0.97	1.78	0.97		1.75	0.97	0.62
14.	Self-injurious	1.45	0.78	1.25	0.65		1.36	0.75	1.38	0.73		1.37	0.74	0.59
15.	Self-esteem	2.03	0.98	1.75	0.93		1.94	0.97	1.91	96.0		1.92	0.97	0.61
19.	Self-reproach	2.45	0.97	1.96	0.98		2.14	1.02	2.36	0.98		2.26	1.00	0.72
20.	Self-deprecation	2.46	1.01	1.92	1.00		2.12	1.05	2.36	1.01		2.25	1.04	0.74
30.	Helplessness	1.93	0.98	1.62	0.91		1.70	0.94	1.90	0.98		1.81	0.97	69.0
Somatic complaints items		17.07	4.20	15.35	4.55	-4.68**	15.71	4.51	16.99	4.24	_3.68**	16.40	4.41	
11.	Somatic complaint	1.97	0.91	1.77	0.89		1.81	0.90	1.97	06.0		1.89	0.91	0.53
17.	Pessimism	2.55	0.94	2.15	1.08		2.31	1.04	2.46	0.99		2.39	1.02	0.55
18.	Fatigue	3.24	0.77	2.97	0.91		3.01	0.91	3.24	9.70		3.14	0.84	0.58
22.	Anger	2.49	0.00	2.27	0.98		2.26	0.97	2.52	0.89		2.40	0.94	0.59
24.	Sleep disturbance	2.16	1.02	2.09	1.04		2.02	1.03	2.23	1.02		2.13	1.03	0.46
27.	Stomachaches	2.09	0.91	1.61	0.80		1.82	0.90	1.98	06:0		1.90	0.90	0.52
28.	Loss of interest	2.58	0.87	2.50	0.99		2.48	0.93	2.60	0.90	:	2.55	0.92	0.50
Total score		62.80	15.62	54.36	14.62	-6.61**	56.98	16.35	61.69	14.95	-4.17**	59.51	15.78	

Total score Corrected item-total correlations for items contra subscale. Note. *p < 0.05, ** $p \le 0.001$, ns. nonsignificant.



Table 3. Standardized factor loadings for the Swedish Reynolds Adolescent Depression Scale second edition for Girls, Boys, Age Groups and Total Sample.

Subscales	Abbreviation	Girls (<i>N</i> = 389) Standardized	Boys ($N = 248$) Standardized	12–15 years ($N = 295$) Standardized	16–20 years ($N = 342$) Standardized	Total sample ($N = 564$) Standardized
Dysphoric m	nood Items					
2.	School anxiety	0.63	0.63	0.65	0.67	0.67
3.	Loneliness	0.71	0.75	0.73	0.71	0.73
6.	Social withdrawal	0.64	0.59	0.66	0.59	0.63
7.	Sadness	0.90	0.89	0.95	0.86	0.92
8.	Crying	0.77	0.77	0.88	0.71	0.81
16.	Irritability	0.74	0.68	0.77	0.69	0.73
21.	Self-pity	0.54	0.71	0.62	0.62	0.63
26.	Worry	0.75	0.84	0.82	0.78	0.80
Anhedonia/r	negative affect items					
1.	Reduced affect	0.87	0.95	0.91	0.87	0.90
5.	Self-worth	0.88	0.77	0.80	0.90	0.85
10.	Anhedonia-peers	0.64	0.56	0.62	0.58	0.60
12.	Low self-worth	0.69	0.66	0.59	0.68	0.63
23.	Reduced speech	0.46	0.35	0.38	0.46	0.41
25.	Anhedonia-general	0.61	0.57	0.55	0.63	0.58
29.	Appetite disturbance	0.60	0.45	0.68	0.48	0.58
Negative sel	f-evaluation items					
4.	Feelings of rejection	0.51	0.44	0.52	0.446	0.485
9.	Worthlessness	0.79	0.80	0.81	0.79	0.80
13.	Discouragement	0.81	0.66	0.80	0.73	0.76
14.	Self-injurious	0.77	0.71	0.78	0.74	0.76
15.	Self-esteem	0.68	0.69	0.72	0.68	0.69
19.	Self-reproach	0.92	0.94	0.93	0.93	0.94
20.	Self-deprecation	0.93	0.96	0.95	0.94	0.95
30.	Helplessness	0.82	0.86	0.82	0.85	0.84
Somatic con	nplaints items					
11.	Somatic complaint	0.68	0.68	0.72	0.65	0.69
17.	Pessimism	0.70	0.75	0.72	0.75	0.73
18.	Fatigue	0.69	0.58	0.67	0.64	0.66
22.	Anger	0.69	0.77	0.72	0.70	0.72
24.	Sleep disturbance	0.56	0.62	0.61	0.52	0.57
27.	Stomachaches	0.61	0.71	0.72	0.64	0.68
28.	Loss of interest	0.60	0.63	0.56	0.64	0.60

Table 4. Individual confirmatory factor analyses across sex (girls, boys) and age group (12-15 years, 16-20 years) for the correlated 4-factor model of Reynolds Adolescent Depression Scale second edition (RADS-2).

Group	n	$\chi^2(df)$	CFI	TLI	RSMEA (90% CI)
Girls	389	11163.444** (399)	0.944	0.939	0.70 [0.066-0.075]
Boys	248	813.538** (399)	0.95	0.945	0.065 [0.058 - 0.071]
12-15 years	295	947.405** (399)	0.953	0.949	0.068 [0.063-0.074]
16-20 years	342	1036.408** (399)	0.943	0.938	0.068 [0.063-0.074]

Note. **p < 0.001. χ^2 : Chi square; df: degrees of freedom; CFI: Comparative Fit Index; RMSEA: Root Mean Square Error of Approximation; CI: Confidence Interval; TLI: Tucker-Lewis Index.

concurrent, convergent and discriminant validity using correlations with other measures of similar and different constructs. These findings are in line with previous findings that self-assessment measures of depression are strongly associated with scores on related internalizing and psychosocial measures such as anxiety, and low self-esteem [51,52]. The confirmatory factor analysis supported the 4-factor structure proposed by Reynolds [15] with acceptable fit indices. Measurement invariance was confirmed using Svetina and Rutkowski guidelines [39] which propose stricter goodness of fit values to establish metric and scalar MI. RADS-2 was found to be invariant across both sex (girls, boys) and age groups (12–15 years, 16–20 years), making it a useful measure allowing interpretation of gender and age differences in the assessment of depression symptoms.

An unexpected finding was the low factor loading on item 23 'Reduced speech' in the anhedonia/negative affect subscale (factor loadings of 0.35 to 0.46). This may have cultural explanations, perhaps in Northern Sweden reduced speech is a normal feature of adolescents and less related to anhedonia/negative affect as compared to in other cultures.

Limitations and strengths

A limitation of this study is that we did not gather data on ethnicity, but only on nationality in general and therefore could not validate the RADS-2 in ethnic minority groups. Furthermore, we relied on self-report measures alone as validation measures for RADS-2. Participants were not geographically stratified and did not match the Swedish general paediatric population. Instead, the participants constituted a convenience sample drawn from four different schools from different socioeconomic areas.

Strengths of the present study are the evaluation of generalizability of scores on this measure over time (i.e. testretest reliability) and the collection of data from different schools from different socioeconomic areas. Another strength is that the robust DWLS estimator (DWLSSS) was used in the confirmatory factor analysis and in the measurement invariance test. The DWLSSS gives more accurate estimates than the DWLS with ordinal data with fewer than 5 categories and hence was used in this study [53]. Furthermore, a majority of validations studies report the DWLS estimator which in this study would have generated better fit indices (CFI = 0.986, TLI = 0.985, RSMEA = 0.069 (90% CI [0.065-0.072]).



Table 5. Measurement Invariance Goodness of Fit for the 4-Factor Model of Reynolds Adolescent Depression Scale second edition (RADS-2) presented by sex and age-group.

Invariance	$\chi^2(df)$	CFI	TLI	RSMEA (90% CI)	Δχ	∆df	Δ CFI	∆ RSMEA
Sex								
Configural	1865.697** (798)	0.951	0.947	0.065 [0.061-0.069]				
Metric	1907.590** (828)	0.950	0.948	0.064 [0.060-0.068]	41.893	30	-0.001	-0.001
Scalar	1908.320** (854)	0.952	0.951	0.062 [0.059-0.066]	0.730	26	0.002	-0.002
Age group								
Configural	1984.697** (798)	0.948	0.944	0.068 [0.065-0.072]				
Metric	2023.640** (828)	0.948	0.945	0.067 [0.064-0.071]	38.943	30	0	-0.001
Scalar	2012.682** (854)	0.949	0.949	0.065 [0.061-0.069]	-10.958	26	0.001	0.002

Note. **p < 0.001. χ^2 : Chi square; df: degrees of freedom; CFI: Comparative Fit Index; RMSEA: Root Mean Square Error of Approximation; CI: Confidence Interval; TLI: Tucker-Lewis Index; Δ CFI: Change in Comparative Fix Index. For sex groups (girls and boys) Satorra-Bentler $\Delta\chi^2$: Config vs. Metric χ^2 (30) = 38.425, p = 0.139. Metric vs. Scaler χ^2 (26) = 35.428, p = 0.103. For age groups: Satorra-Bentler $\Delta\chi^2$: Config vs. Metric χ^2 (30) = 26.83, p = 0.6322. Metric vs. Scaler χ^2 (26) = 22.736, p = 0.648.

Table 6. Pearson's correlations between Swedish RADS-2 subscales and cross-validity instruments: Beck Youth Inventories of Emotional and Social Impairment subscales Depression (BYI-D); and Anger (BYI-A); the World Health Organization Wellness Index (WHO-5); the Patient Reported Outcome Measurements Information System (PROMIS) subscales Anxiety (PROMIS Anxiety); and Peer relationship (PROMIS Friend).

Measure	Dysphoric mood	Anhedonia/negative affect	Negative self-evaluation	Somatic complaints	RADS-2 total scale
RADS-2 Dysphoric mood		0.36**	0.77**	0.76**	0.90**
RADS-2 Anhedonia/negative affect			0.55**	0.35**	0.63**
RADS-2 Negative self-evaluation				0.69**	0.92**
RADS-2 Somatic complaints					0.85**
BYI-D .	0.76**	0.56**	0.85**	0.75**	0.88**
PROMIS Anxiety	0.67**	0.43**	0.64**	0.57**	0.70**
BYI-A	0.61**	0.46**	0.66**	0.68**	0.73**
WHO-5	0.62**	0.48**	0.64**	0.65**	0.72**
PROMIS Frnd	-0.38**	-0.50**	-0.44**	-0.33**	-0.48**

^{**}p < 0.01 (2-tailed).

We therefore conclude that since Reynolds Adolescent Depression Scale second edition is an established instrument compatible with DSM and other diagnostic criteria such as the ICD and Research Diagnostic Criteria (RDC) and thus in line with theory, the fit for the 4-factor model is considered to be good.

The RADS-2 is extensively used throughout the world both clinically and in research settings and using a validated Swedish version of RADS-2 as an outcome in Swedish research studies would facilitate comparison with international findings. Since RADS-2 measures different dimensions of MDD, is in line with the RDoC initiative and still is in coherence with the DSM system which currently is the 'gold standard' for diagnostic classification used in Sweden, it creates a useful bridge between these two classification systems.

Our aim was to validate the scale in a normative sample and we found evidence that the Swedish version of RADS-2 had good concurrent, convergent and discriminant validity and excellent reliability for the RADS-2 total scale. Goodness of fit for the established 4-factor structure was good and the scale was invariant across both sex (girls, boys) and age groups (12–15 years, 16–20 years). The next steps are to validate the RADS2 in a clinical sample, to evaluate the translation with differential item functioning analysis comparing our sample with an English speaking sample and to report Swedish norms in accordance with the International Test Commission guidelines [54].

In summary, the RADS-2 is a 'reliable and useful instrument' [22] also in a Swedish community context.

Acknowledgments

The authors thank the students who participated in the study. Thanks also to the Queen Silvia Jubilee Fund.

Disclosure statement

The authors report no conflicts of interest.

Notes on contributors

Ida Blomqvist is a PhD student at the Department of Clinical Science, the Child- and Adolescent Psychiatry unit, Umeå University and a resident at the Child and Adolescent Psychiatry Outpatient Clinic in Örnsköldsvik.

Erik Ekbäck is a PhD student at the Department of Clinical Science, the Child- and Adolescent Psychiatry unit, Umeå University. He is also a general practitioner resident in Umeå and Örnsköldsvik.

Inga Dennhag, PhD, psychologist and psychotherapist. She teaches and work with research at the Child and Adolescent Psychiatry at Umeå University in Sweden. Her main research areas are psychometrics, teenage depression and trauma. Year 2016, she received the title "Excellent teacher". Year 2017, the book *Power and Psychotherapy* came out.

Eva Henje is a senior specialist and associate professor in Child and Adolescent Psychiatry at Department of Clinical Science at Umeå University. Her clinical work and research is mainly focused on trauma-and stress related disorder and depression in youth.

ORCID

Ida Blomqvist (b) http://orcid.org/0000-0003-0487-5552



Erik Ekbäck (D) http://orcid.org/0000-0003-0496-7158 Inga Dennhag (b) http://orcid.org/0000-0001-7336-1657 Eva Henje (i) http://orcid.org/0000-0001-5790-0518

References

- Ferrari AJ, Charlson FJ, Norman RE, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. PLoS Med. 2013;10:e1001547.
- Mathers C, Fat DM, Boerma JT. The global burden of disease: 2004 update. Geneva (Switzerland): World Health Organization;
- Birmaher B, Ryan ND, Williamson DE, et al. Childhood and adolescent depression: a review of the past 10 years. Part I. J Am Acad Child Adolesc Psychiatry. 1996;35:1427-1439.
- Korczak DJ, Goldstein BI. Childhood onset major depressive disorder: course of illness and psychiatric comorbidity in a community sample. J Pediatr. 2009;155:118-123.
- Thapar A, Collishaw S, Pine DS, et al. Depression in adolescence. Lancet. 2012;379:1056-1067.
- Joynt KE, Whellan DJ, O'Connor CM. Depression and cardiovascu-[6] lar disease: mechanisms of interaction. Biol Psychiatry. 2003;54:
- Spiegel D, Giese-Davis J. Depression and cancer: mechanisms and disease progression. Biol Psychiatry. 2003;54:269-282.
- Wulsin LR, Vaillant GE, Wells VE. A systematic review of the mortality of depression. Psychosom Med. 1999;61:6-17.
- Alexandrino-Silva C, Wang YP, Carmen Viana M, et al. Gender differences in symptomatic profiles of depression: results from the São Paulo Megacity Mental Health Survey. J Affect Disord. 2013; 147:355-364.
- [10] Blom EH, Forsman M, Yang TT, et al. Latent classes of symptoms related to clinically depressed mood in adolescents. Scand J Child Adolesc Psychiatr Psychol. 2014;2:19-28.
- Costello EJ, Mustillo S, Erkanli A, et al. Prevalence and development of psychiatric disorders in childhood and adolescence. Arch Gen Psychiatry. 2003;60:837-844.
- [12] Ford T, Goodman R, Meltzer H. The British Child and Adolescent Mental Health Survey 1999: the prevalence of DSM-IV Disorders. J Am Acad Child Adolesc Psychiatry. 2003;42:1203-1211.
- [13] Schuch JJJ, Roest AM, Nolen WA, et al. Gender differences in major depressive disorder: results from the Netherlands study of depression and anxiety. J Affect Disord. 2014;156:156-163.
- [14] Insel T, Cuthbert B, Garvey M, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry. 2010;167:748-751.
- [15] Reynolds WM. 2002. Reynolds Adolescent Depression Scale: professional manual. 2nd ed. Odessa (FL): Psychological Assessment Resources.
- Fonseca-Pedrero E, Wells C, Paino M, et al. Measurement invari-[16] ance of the Reynolds Depression Adolescent Scale across gender and age. Int J Test. 2010;10:133-148.
- [17] Henje Blom E, Tymofiyeva O, Chesney MA, et al. Feasibility and preliminary efficacy of a novel RDoC-based treatment program for adolescent depression: "Training for Awareness Resilience and Action" (TARA) - a pilot study. Front Psychiatry. 2017;7:208.
- [18] Hyun MS, Nam KA, Kang HS, et al. Reynolds Adolescent Depression Scale - Second Edition: initial validation of the Korean version. J Adv Nurs. 2009;65:642-651.
- [19] King CA, O'Mara RM, Hayward CN, et al. Adolescent suicide risk screening in the emergency department. Acad Emerg Med. 2009; 16:1234-1241.
- [20] Stockings E, Degenhardt L, Lee YY, et al. Symptom screening scales for detecting major depressive disorder in children and adolescents: a systematic review and meta-analysis of reliability, validity and diagnostic utility. J Affect Disord. 2015;174:447-463.
- Walker L, Merry S, Watson PD, et al. The Reynolds Adolescent Depression Scale in New Zealand adolescents. Aust N Z J Psychiatry. 2005;39:136-140.

- Osman A, Gutierrez PM, Bagge CL, et al. Reynolds adolescent depression scale-second edition: a reliable and useful instrument. J Clin Psychol. 2010;66:1324-1345.
- Sund AM, Larsson B, Wichstrøm L. Depressive symptoms among young Norwegian adolescents as measured by The Mood and Feelings Questionnaire (MFQ). Eur Child Adolesc Psychiatry. 2001; 10:222-229.
- Sweden S. 1984. Swedish socio-economic classification (SEI). Available from: https://www.scb.se/contentassets/22544e89c6f34 ce 7ac2e6fefbda407ef/english_ov9999_1982a01_br_x11op8204-3. pdf
- [25] Sweden S. 2020. SEI, yrkesförteckning version 2020-04-16. Available from: https://www.scb.se/contentassets/22544e89c6f 34ce7ac2e6fefbda407ef/sei index webb 20200416.pdf
- Beaton ED, Bombardier EC, Guillemin EF, et al. Guidelines for the [26] Process of Cross-Cultural Adaptation of Self-Report Measures. Spine (Phila Pa 1976). 2000;25:3186-3191.
- [27] Henje E, Rindestig FC, Gilbert P, et al. Psychometric validity of the Compassionate Engagement and Action Scale for Adolescents: a Swedish version. Scand J Child Adolesc Psychiatry Psychol. 2020; 8:70-80.
- [28] Beck JS, Beck AT, Jolly J. 2001. Manual for the Beck Youth Inventories of Emotional and Social Impairment. San Antonio (TX): The Psychological Corporation.
- [29] Dunerfeldt ME, Anna Söderström B. 2010. Bedömningsinstrument inom BUP i Stockholm Kartläggning och faktasammanställning. Available from: http://www.bup.se/globalassets/dokument/ombup/publicerat/aktuella rapporter/rapport bedomningsinstrument inom bup i stockholm.pdf
- [30] SBU. Diagnostik och uppföljning av förstämningssyndrom. En systematisk litteraturöversikt. Stockholm: Statens beredning för medicinsk utvärdering (SBU); 2012. SBU-rapport nr 212.
- [31] Topp CW, Østergaard SD, Søndergaard S, et al. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom. 2015;84:167-176.
- [32] Blom EH, Bech P, Hogberg G, et al. Screening for depressed mood in an adolescent psychiatric context by brief self-assessment scales-testing psychometric validity of WHO-5 and BDI-6 indices by latent trait analyses. Health Qual Life Outcomes. 2012;
- [33] Irwin D, Stucky B, Langer M, et al. An item response analysis of the pediatric PROMIS anxiety and depressive symptoms scales. Oual Life Res. 2010;19:595-607.
- [34] Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. Psychol Bull. 1980;88:
- [35] Hooper D, Coughlan J, Mullen MR. Structural equation modelling: guidelines for determining model fit. Electron J Bus Res Methods. 2008:6:53-60.
- [36] Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Modeling. 1999;6:1-55.
- Li CH. Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. Behav Res Methods. 2016a;48:936-949.
- [38] Li CH. The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. Psychol Methods. 2016b;21:369-387.
- [39] Svetina D, Rutkowski L, Rutkowski D. Multiple-group invariance with categorical outcomes using updated guidelines: an illustration using Mplus and the lavaan/semTools packages. Struct Equ Modeling. 2019;27:1-20.
- Kite BA, Jorgensen TD, Chen PY. Random permutation testing [40] applied to measurement invariance testing with ordered-categorical indicators. Struct Equ Modeling. 2018;25:573-587.
- [41] Rutkowski L, Svetina D. Measurement invariance in international surveys: categorical indicators and fit measure performance. Appl Meas Educ. 2017;30:39-51.

- [42] Svetina D, Rutkowski L. Multidimensional measurement invariance in an international context: fit measure performance with many groups. J Cross Cult Psychol. 2017;48:991-1008.
- McNeish D. Thanks coefficient alpha, we'll take it from here. [43] Psychol Methods. 2018;23:412-433.
- Raykov T, Marcoulides GA. Thanks coefficient alpha, we still need [44] you! Educ Psychol Meas. 2019;79:200-210.
- [45] Nunnally JC. 1994. Psychometric theory. New York: McGraw-Hill.
- Kennedy DM, Stratford PW, Wessel J, et al. Assessing stability and [46] change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty. BMC Musculoskelet Disord. 2005;6:3.
- Weir JP. Quantifying test-retest reliability using the intraclass cor-[47] relation coefficient and the SEM. J Strength Cond Res. 2005;19: 231-240.
- [48] Cohen J. 2013. Statistical power analysis for the behavioral sciences. New York (NY): Academic Press.

- [49] Team, RDC. 2019. R: a language and environment for statistical computing. Vienna (Austria). Available from: http://www.R-project.org/
- [50] Rosseel Y. lavaan: an R package for structural equation modeling. J Stat Softw. 2012;48:1-36.
- [51] Esposito C, Clum G. Specificity of depressive symptoms and suicidality in a juvenile delinquent population. J Psychopathol Behav Assess. 1999;21:171-182.
- [52] Smith M, Calam R, Bolton C. Psychological factors linked to selfreported depression symptoms in late adolescence. Behav Cogn Psychother. 2009;37:73-85.
- Rhemtulla M, Brosseau-Liard PÉ, Savalei V. When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. Psychol Methods. 2012;17: 354-373.
- [54] ITC. ITC Guidelines for Translating and Adapting Tests (Second Edition). Int J Test. 2017;18:101-134.