

Nordic Journal of Psychiatry



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

Psychometric properties of the Swedish version of the Reynolds Adolescent Depression Scale second edition (RADS-2) in a clinical sample

Erik Ekbäck, Ida Blomqvist, Inga Dennhag & Eva Henje

To cite this article: Erik Ekbäck, Ida Blomqvist, Inga Dennhag & Eva Henje (2023) Psychometric properties of the Swedish version of the Reynolds Adolescent Depression Scale second edition (RADS-2) in a clinical sample, Nordic Journal of Psychiatry, 77:4, 383-392, DOI: 10.1080/08039488.2022.2128409

To link to this article: https://doi.org/10.1080/08039488.2022.2128409

| 9 | © 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. | Published online: 04 Nov 2022. |
|---|---|--------------------------------|
| | Submit your article to this journal $oldsymbol{G}$ | Article views: 590 |
| Q | View related articles 🗹 | Usew Crossmark data ☑ |
| 4 | Citing articles: 1 View citing articles 🗹 | |

Tavlor & Francis Taylor & Francis Group

ARTICLE



Psychometric properties of the Swedish version of the Reynolds Adolescent Depression Scale second edition (RADS-2) in a clinical sample

Erik Ekbäck (10), Ida Blomqvist (10), Inga Dennhag (10) and Eva Henje (10)

Department of Clinical Science, Umeå University, Umeå, Sweden

ABSTRACT

Objective: Observed and predicted increases in the global burden of disease caused by major depressive disorder (MDD) highlight the need for psychometrically robust multi-dimensional measures to use for clinical and research purposes. Reynolds Adolescent Depression Scale second edition (RADS-2) is an internationally well-validated scale measuring different dimensions of adolescent depression. The Swedish version has previously only been evaluated in a normative sample.

Methods: We collected data from patients in child and adolescent psychiatry and primary care and performed: (1) Confirmatory factor analysis (CFA) to evaluate the established four-factor structure, (2) Analyses of reliability and measurement invariance, (3) Analyses of convergent and discriminant validity using the Montgomery-Asberg Depression Rating Scale, the depression subscales of the Beck Youth Inventories and the Revised Child Anxiety and Depression Scale, as well as the Patient Reported Outcome Measurements Information System, peer-relationships and physical activity item banks.

Results: Recruited participants (n = 536, 129 male and 407 female, mean age 16.45 years, SD = 2.47, range 12 – 22 years) had a variety of psychiatric diagnoses. We found support for the four-factor structure and acceptable to good reliability for the subscale and total scores. Convergent and discriminant validity were good. Measurement invariance was demonstrated for age, sex, and between the present sample and a previously published normative sample. The RADS-2-scores were significantly higher in the present sample than in the normative sample. In this clinical study, the Swedish RADS-2 demonstrated good validity and acceptable to good reliability. Our findings support the use of RADS-2 in Swedish clinical and research contexts.

ARTICLE HISTORY

Received 26 May 2022 Revised 22 August 2022 Accepted 21 September 2022

KEYWORDS

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

Introduction

It is predicted that major depressive disorder (MDD) will soon top the list of mental and physical disorders with the largest negative impact on global health [1], and it is already causing substantial disability worldwide [2]. The teenage years is a vulnerable developmental period during which the MDD prevalence is increasing, particularly in females [3]. This early onset of MDD increases the risk for recurrent depressive episodes [4] and increases all-cause mortality as well as suicide rates [5]. The clinical picture of teenage MDD differs from that of adults. For example, the Diagnostic and Statistical Manual of Mental disorders, fifth edition (DSM-5) allows for the replacement of depressed mood with irritability as a core diagnostic symptom criterion [6]. At the same time, both pathophysiological and neurodevelopmental similarities are shared between teenagers and adults up to their mid-twenties [7,8], which has warranted research spanning across this critical age-range [9]. In the evaluation of symptom severity and treatment effectiveness for individual patients, as well as more broadly for research with national and international ambitions, measures of depression that are

valid across different contexts, age-groups, and languages are clearly needed.

The Reynolds Adolescent Depression Scale second edition (RADS-2) is one such internationally well-validated ageappropriate measure, and by means of self-report it quantifies four dimensions of depression: dysphoric mood, anhedonia/negative affect, negative self-evaluation, and somatic complaints [10]. RADS-2 is compatible with both the DSM-5 and the International Classification of Disease, eleventh edition (ICD-11), and the scale is widely used both clinically as well as in research [11-19]. The four-factor structure of RADS-2 has been supported in several confirmatory factor analyses, see, e.g. [11,13,16], and convergent as well as discriminant validity are good to excellent both in non-clinical [11,13] and in clinical [17] samples. Reliability has also been demonstrated in large samples [10]. We have translated RADS-2 to Swedish and have replicated these findings in a normative sample study previously published in this journal [20]. In that sample measurement, invariance was also confirmed for sex and age-group [20]. Since the scale is intended for clinical application [10] and since valid outcome-measures are needed for both ongoing and future clinical trials, including in participants from a wider age-

Table 1. Descriptive statistics of the sample.

| Years of age 12 13 14 | 20 (3.73) 34 (6.34) 55 (10.26) |
|--|--------------------------------------|
| 13 | 34 (6.34) |
| | |
| 14 | 55 (10.26) |
| | |
| 15 | 89 (16.60) |
| 16 | 96 (17.91) |
| 17 | 112 (20.90) |
| 18 | 38 (7.09) |
| 19 | 20 (3.73) |
| 20 | 17 (3.17) |
| 21 | 24 (4.48) |
| 22 | 31 (5.78) |
| Total | 536 (100) |
| Sex | () |
| Male | 129 (24.07) |
| Female | 407 (75.93) |
| Total | 536 (100) |
| Recruited from | , , , , |
| Child and adolescent psychiatry | 382 (71.27) |
| Youth clinic | 56 (10.45) |
| Health-care center | 3 (0.56) |
| Recruitment route not registered | 95 (17.72) |
| Total | 536 (100) |
| Concurrent primary ICD-diagnosis category | ` , |
| Diagnosis not extracted | 249 (46.5) |
| MDD or other affective disorder | 81 (15.1) |
| Anxiety-disorder | 44 (8.2) |
| Stress and trauma-related disorders, including | 5 (0.9) |
| PTSD and complex PTSD | , |
| Anorexia/bulimia nervosa or other eating disorder | 15 (2.8) |
| OCD | 10 (1.9) |
| Autism spectrum disorder | 13 (2.4) |
| ADHD/ADD conduct disorder | 107 (20.0) |
| Psychotic disorder | 1 (0.2) |
| Tics or Tourettes syndrome | 2 (0.4) |
| Gender dysphoria or dysmorphophobia | 9 (1.7) |
| Total | 536 (100) |
| Socioeconomic classification of parent(s)/caregiver(s) | 330 (100) |
| Workers | 106 (19.78) |
| Assistant and intermediate- | 100 (15.70) |
| Non-manual workers | 186 (34.70) |
| Professionals, civil servants, and executives | 130 (24.25) |
| Self-employed | 48 (8.96) |
| Unknown | 66 (12.31) |
| Total | 536 (100) |

Note. MDD: Major Depressive Disorder; OCD: Obsessive Compulsive Disorder; PTSD: Post Traumatic Stress Disorder: ADHD: Attention Deficit Hyperactivity Disorder; ADD: Attention Deficit Disorder.

range than previously studied [9], we here present data on the psychometric properties of RADS-2 in a heterogenous clinical sample with affective symptomatology.

In this study, we aimed to evaluate the RADS-2 factor structure, validity, and reliability as well as measurement invariance to determine whether the scale measures the construct equivalently in males compared to females, and in teenagers 12–17 years old compared to young adults 18–22 years old. We also aimed to compare individual scores from a normative non-clinical sample with scores from this clinical sample to test the hypothesis that the scores would be higher in the clinical sample. To support such comparisons, we also aimed to test measurement invariance for the clinical and non-clinical samples.

Materials and methods

This cross-sectional study was approved by the regional ethical review board at Umeå University in Sweden (D.nr 2018/

59-31), by PAR-inc., the publisher and copyright holder of RADS-2, as well as by the manager of each participating clinic. Written informed consent was collected from all participants before inclusion. Additional parental consent was collected for participants below 15 years of age. reimbursement equivalent to 20 Euro was provided to the participants.

Participant recruitment and data collection

Participants were recruited from four child and adolescent psychiatry clinics, one primary care youth-clinic and one primary care health-clinic; all in four Swedish cities/towns with population-range 8000 – 130,000. Flyers were posted in the waiting-rooms of the clinics. Information was also sent out by mail or SMS to patients with affective disorders who were admitted to the clinics and either waitlisted for treatment or in active treatment, for details see the eligibility criteria below. Those who did not respond to these invitations were contacted over the phone once.

Eligibility criteria were: (1) Being between 12 and 22 years of age, (2) Being a patient at any of the recruiting sites, (3) Having self-reported or parent-reported symptoms of depression and/or anxiety (all comorbidities were allowed), (4) For individuals with a recent history of suicide attempt or psychiatric inpatient-care a minimum time of three months had to have passed from the suicidal event or since discharge from hospitalization, and (5) Fluency in written Swedish and ability to complete the questionnaires.

Eligible participants were sent a link by email to an online platform where they signed in from their preferred location and device, provided written informed consent, and responded to the guestionnaires. The order of the scales in the questionnaire was altered between participants to prevent a bias effect of fatigue in replying. At the end of questionnaire-data collection additional data on psychiatric diagnoses that had been given one month before and after the self-rating was extracted from the individual medical records of the participants that were recruited from child and adolescent psychiatry. Descriptive statistics of the participant demographics are presented in Table 1. Data collection was performed between 2019 and 2022.

We also re-analyzed data from a previously published normative sample of n = 637 [20], to evaluate differences in RADS-2 scores between the normative and clinical samples. The methods used for participant recruitment and data collection for the normative sample have been described in detail in our previous publication [20].

Instruments

Reynolds adolescent depression scale second edition (RADS-2)

RADS-2 is a 30-item self-rating scale with brief self-statements like 'I feel like crying' [10]. Response options are ordinal on a four-point scale ranging from 'almost never' to 'most of the time'. The four subscales are 'dysphoric mood', 'anhedonia/negative affect', 'negative self-evaluation', and



'somatic complaints'. Items in the anhedonia/negative affect subscale are reversely phrased, e.g. 'I feel happy' and hence reversely coded. The scale sum raw score ranges from 30 to 120 and higher scores indicate more severe symptomatology [10]. RADS-2 has been translated to Swedish and previously validated in a normative sample [20].

Beck vouth inventories

From the Beck Youth Inventories of Emotional and Social Impairment [21] we specifically used the depression subscale (BYI-D). BYI-D consists of 20 brief self-statement questions like, e.g. 'I feel sad', with responses on a four-point ordinal scale ranging from 'never' to 'always'. The sum raw score range is 0-60 and higher scores indicate more severe symptomatology. Internationally high internal consistency as well as test-retest reliability have been demonstrated in large samples [21]. Even though the discriminative ability of the scale has been questioned [22] findings of good convergent validity have been replicated in a clinical adolescent outpatient sample [23]. In Sweden BYI-D is widely used in Child and Adolescent Psychiatry to evaluate depression severity [24]. It is also recommended by the Swedish Agency for Health Technology Assessment and Assessment of Social Services in screening for MDD [25]. In the current sample the BYI-D Cronbach's alpha was 0.95 (95% CI 0.94-0.95).

Montgomery-Asberg depression rating scale (MADRS)

The MADRS is a scale that is widely used to assess depression severity [26–28]. We used the self-rating version of the scale, which includes nine items on a seven-point ordinal scale, e.g. reported sadness, with a sum raw score range of 0-54. Higher scores indicate more severe symptomatology [29]. Reliability and validity are good in Swedish adolescent psychiatric outpatients [29]. In this sample, the MADRS Cronbach's alpha was 0.89 (95% CI 0.87-0.90).

Revised Child Anxiety and Depression Scale (RCADS)

From the RCADS we specifically used the depression subscale (RCADS-depression), which consists of 10 items rating the extent to which one is, e.g. 'feeling sad or empty', on a four-point ordinal scale from 'never' to 'always' [30]. The scale has good validity and reliability in clinical populations of children and adolescents in different assessment settings, countries, and languages, see, e.g. [31-34]. The sum raw score range is 0-30 with higher scores indicating more severe symptomatology [31]. In this sample, the RCADSdepression Cronbach's alpha was 0.90 (95% CI 0.89-0.91).

Patient-reported informaoutcome measurement tion system

The National Institutes of Health developed the Patient Outcome Measurement Information System (PROMIS), which contains item banks for various health and lifestyle dimensions [35]. We used the PROMIS Pediatric Bank version 1.0 [36] - Physical activity (PROMIS-physical activity) and PROMIS Pediatric Bank version 2.0 [36] Peer-relationships (PROMIS-peer-relationships) in this study. These item banks consist of 10 and 15 questions, respectively, each framed in past tense starting 'In the last seven days ... ', e.g. ' ... I was able to count on my friends.' Responses are recorded on a five-point ordinal scale ranging from 'Never' to 'Almost always'. The sum raw score range is 0-40 (PROMIS-physical activity) and 0-60 (PROMIS-peer-relationships), and higher scores indicate more of the measured construct. For more information on item definitions and the concepts behind them, see [36]. The item banks used for this study have been translated and culturally adapted for Swedish adolescents [37] and the former of the two has been psychometrically evaluated in Swedish adolescents with good reliability [38]. In this sample, the Cronbach's alphas were 0.93 (95% CI 0.92-0.94) and 0.90 (95% CI 0.88-0.91) for PROMIS-physical activity and PROMIS-peer-relationships, respectively.

Data analysis

We used standard measures for descriptive statistics. On RADS-2 the missing item-level data range was 2–5 (0.4–0.9%) for each individual item. Thirteen participants (2.4%) had missing values on any RADS-2 item and Little's test was not significant (Chi-square (χ^2) 264.49, DF 284, p = 0.79). We assumed the missingness mechanism was completely random and removed these individuals from dataset altogether. In a similar way for validity-analyses listwise deletion of individuals with missing item-level data was applied on BYI-D, MADRS, RCADS-depression, and PROMIS item banks to allow for total-score calculations (the range of missing item-level data was 0-6, i.e. 0-1.1% for individual items). Total sum-scores of items on ordinal scales were conservatively treated as ordinal variables throughout the analysis. To analyze sex- and age-group (12-17 years and 18-22 years) differences in RADS-2 subscale scores, and totalscale scores we used the Mann-Whitney U test, which was also used to compare mean RADS-2 scores in the clinical and normative samples.

To test the four-factor model previously proposed [10,11,13,16] we performed a four-factor correlated model confirmatory factor analysis (CFA). χ^2 tests were performed, although sensitive to sample size [39], and the comparative fit index (CFI), Tucker-Lewis' index (TLI), mean square error approximation (RMSEA), and standardized root mean residual (SRMR) were used to evaluate the goodness-of-fit. The acceptability of model fit was evaluated using the following criteria: (1) CFI of 0.90-0.94, (2) TLI of 0.80-0.89, and (3) RMSEA of < 0.06 (95% CI 0.00 - 0.08) [39–41] and SRMR <0.08 [41]. We used the robust scaled diagonally weighted least squares estimator [42,43].

McDonald's coefficient Omega [44] was used to test reliability and the following well-established cut-offs were used interpretation of internal consistency: 0.7 = acceptable, > 0.8 = good, and > 0.9 = excellent [45].

Measurement invariance/equivalence was tested in a specific forward procedure for ordered variables following the model identification approach of Wu and Estabrook [46] and as laid out in detailed guidelines by Svetina et al. [47]. This was done to test hypotheses of RADS-2 being understood and measured equivalently in males and females, in the different age-groups (12-17 and 18-22), and in the clinical and non-clinical samples respectively. The following consecutive steps were performed: (1) Configural invariance tests to evaluate the factor structure, with separate individual CFAs on males/females and in both age-groups (12-17 and 18-22). The normative sample CFA has been reported elsewhere [20]. (2) Threshold invariance tests to evaluate the equivalence of thresholds. (3) Metric invariance tests to evaluate the equivalence of thresholds and factor loadings. (4) Strong/scalar invariance tests to evaluate the equivalence of thresholds, factor loadings, and intercepts.

By sequentially performing models with increasingly stringent constraints in this way, and by comparing each model to the previous one; invariance achieved at the scalar level indicates that the scores are not influenced by item-level group differences and that latent means are comparable across groups. In this context, the χ^2 test has high power and inflated Type I error rates have been observed [47]. Therefore, to determine whether measurement invariance had been achieved at a specific level the following cut-offs for change in fit index was considered: $\Delta CFI = -0.002$ and Δ RMSEA = 0.05 for thresholds and Δ CFI = -0.002 and Δ RMSEA = 0.01 for thresholds and factor loadings [47]. We also considered the Satorra–Bentler scaled χ^2 difference test statistic; and non-significant p values were interpreted as indicative of model equivalence [48]. This invariance testing procedure is strong and required for latent means to be compared across groups [47].

To test convergent and discriminant validity we performed Spearman's correlations between RADS-2 and established measures of depression (BYI-D, MADRS, and RCADSdepression) as well as between RADS-2 and constructs that are theoretically distinct from depression (the PROMIS-item banks specified above). Correlations between 0.1 and 0.29 were interpreted as small, 0.3-0.49 as medium, and 0.50 and above as large [49].

All analyses were two-tailed, statistical uncertainties are presented in 95% confidence intervals, except for RMSEA where a 90% interval is current standard, and a significance level of 0.05 was used. We analyzed data using SPSS statistics version 28 (IBM Corp., Armonk, NY) and R [50]. The structural equation modeling used for CFA and measurement invariance modeling was performed in R using the Lavaan package version 0.6-3 [51].

Results

Descriptive statistics

The percentage of invited participants who did not respond or declined to participate was 75%. Descriptive statistics of the sample are presented in Table 1. Mean age was 16.45 years (SD = 2.47), 95.10% were born in Sweden, and 80.10% were living with one or both parents. In Table 1, the participants' primary ICD-diagnosis at the time of data collection is reported as well as their households' socioeconomic status according to the classification system used by Statistics Sweden [52,53].

Means and standard deviations for all RADS-2 items, subscale scores and total scores are reported by age-group (12–17 years and 18–22 years) and sex (male/female), as well as for the whole sample in Table 2. Corrected item to total subscale-score correlations are also shown in Table 2.

Mean RADS-2 total score for the whole clinical sample 76.30 (SD = 18.26), median 78.00 (IQR = 26.00). Significant sex differences were found, with females scoring higher than males on all subscales and the total scale, see Table 2 for details. The only significant age-difference was that the 12-17-year-olds scored higher than the 18-22-yearolds on the Anhedonia/Negative affect subscale, see Table 2 for details. In the normative sample, the RADS-2 total score was significantly lower than in the clinical sample, with normative sample mean 59.61 (SD = 15.79), median 58.00, total n = 1124, Mann-Whitney U = 237663.00, p < 0.001. Each of the four individual subscale scores was also lower in the non-clinical sample, U-value range 236865.00 - 244185.50, all at p < 0.001.

Factor structure

Standardized factor loadings for all RADS-2 items are presented in Table 3, both by age-group (12-17 and 18-22) and sex categories, and for the whole sample. In the whole sample standardized factor loading range was 0.34 (item 21) to 0.96 (item 20).

The test for the whole sample model fit for the four-factor structure was significant (χ^2 (399) = 1946.54, p < 0.001), and fit indices were as follows: CFI = 0.933, TLI = 0.927, RMSEA = 0.085 (90% CI 0.081-0.089), and SRMR 0.068.

Reliability and measurement invariance

Reliability measures for all subscales as well as for the total scale were acceptable to good, see Table 4.

Individual CFAs in the different age and sex-groups are presented in Table 5. Configural CFAs in the age, sex, and clinical/non-clinical sample groups resulted in acceptable fit indices, with the exception of RMSEA, supporting configural invariance. Invariance with regards to (1) thresholds, (2) thresholds and factor loadings, and (3) thresholds, factor loadings, and intercepts was also demonstrated between age and sex-groups, as well as between the clinical and non-clinical samples (although some scaled χ^2 difference tests were significant and one Δ CFI value touched the specified cut-off), see Table 5 for details.

Validity

Spearman's correlation coefficients between RADS-2 total scale and RADS-2 subscales, as well as between RADS-2 total scale and validation-instruments are shown in Table 6. The internal correlations of RADS-2 ranged from (Anhedonia/negative affect and Somatic complaints subscales) to 0.94 (Negative self-evaluation subscale and total

Table 2. Means and standard deviations for RADS-2 items, subscale scores and total scores by age-group, sex, and for the whole sample.

| | | 12–17 years | | (N = 130) | 30) | 7-50079 | (N = 407) | (N = 407) | (N = 129) | 129) | 7-50079 | whole sample $(N=536)$ | ample (36) | ltem to |
|-----------------------------------|------------------------------|-------------|-------|-----------|-------|----------------------|--------------|-----------|---------------|-------|---------|------------------------|---------------|---------------|
| Subscales and items | Abbreviation | M M | SD | N | SD | 2-3006 | W | SD | × | SD | 2,30,6 | M | SD | correlation |
| Dysphoric mood items | | 21.83 | 5.56 | 22.79 | 4.31 | —1.38 ns | 23.38 | 4.46 | 17.90 | 5.63 | -9.37** | 22.06 | 5.31 | |
| 2. | School anxiety Loneliness | 3.08 | 0.98 | 3.15 | 0.93 | | 3.27 | 0.90 | 2.56 | 0.98 | | 3.10 | 0.97 | **99:0 *** |
| 9. | Social withdrawal | 2.76 | 1.10 | 2.85 | 1.04 | | 2.95 | 1.00 | 2.25 | 1.67 | | 2.78 | 1.09 | 0.70 |
| 7. | Sadness | 3.00 | 0.90 | 3.07 | 9.70 | | 3.20 | 0.74 | 2.43 | 0.98 | | 3.02 | 0.87 | 0.77** |
| œ. | Crying | 2.57 | 0.97 | 2.62 | 0.82 | | 2.80 | 0.84 | 1.87 | 0.89 | | 2.58 | 0.94 | 0.72** |
| 16. | Irritability | 2.77 | 0.94 | 2.64 | 0.80 | | 2.88 | 0.84 | 2.29 | 0.97 | | 2.74 | 0.91 | 0.64** |
| 21. عر | Self-pity | 1.96 | 0.94 | 2.77 | 0.93 | | 2.09 | 0.96 | 1.8.1 | 1.07 | | 2.02 | 0.94 | 0.48 |
| zo. Anhedonia/negative affect | , io | 14.92 | 4.97 | 13.58 | 4.12 | -2.67* | 15.36 | 4.77 | 2.30 12.21 | 4.13 | -6.57** | 14.60 | 4.81 | 0.72 |
| items | | | | | | | | | | | | | | |
| 1. | Reduced affect | 2.13 | 96:0 | 1.99 | 0.79 | | 2.23 | 0.92 | 1.66 | 0.77 | | 2.10 | 0.92 | 0.75 |
| 5. | Self-worth | 2.55 | 1.06 | 2.38 | 0.93 | | 2.64 | 1.00 | 2.11 | 1.02 | | 2.51 | 1.03 | 0.75 |
| 10. | Anhedonia-peers | 1.95 | 1.04 | 1.59 | 0.87 | | 1.97 | 1.04 | 1.53 | 0.87 | | 1.86 | 1.01 | 0.71** |
| 12. | Low self-worth | 1.99 | 0.97 | 1.79 | 0.87 | | 2.00 | 0.95 | 1.76 | 0.92 | | 1.94 | 0.95 | 0.68** |
| 23. Jr | Reduced speech | 2.45 | 1.12 | 2.20 | 0.98 | | 2.46 | N.08 | 2.15 | 1.13 | | 2.39 | 90.1 | 0.70 |
| 23. 29 | Annetite disturbance | 04 | 100 | 1.73 | 0.70 | | 5.1 7.1.7 | 0.00 | 153 | 0.7.5 | | 1 98 | 0.00 | 0.75 |
| Negative self-evaluation Items | | 19.75 | 69.9 | 19.98 | 5.57 | 0.16 ns | 21.27 | 5.99 | 15.22 | 5.57 | -9.17** | 19.81 | 6.43 | |
| 4. | Feelings of rejection | 1.77 | 1.00 | 1.69 | 0.92 | | 1.82 | 1.00 | 1.53 | 0.88 | | 1.75 | 0.98 | 0.56 |
| 9. | Worthlessness | 2.37 | 1.04 | 2.48 | 1.05 | | 2.57 | 1.01 | 1.83 | 0.91 | | 2.40 | 1.04 | 0.79 |
| 13. | Discouragement | 2.52 | 1.11 | 2.72 | 1.06 | | 2.77 | 1.05 | 1.94 | 1.04 | | 2.57 | 1.10 | 0.78** |
| 14. | Self-injurious | 2.16 | 1.15 | 2.03 | 1.05 | | 2.33 | 1.12 | 1.50 | 0.88 | | 2.13 | 1.13 | 0.77 |
| 15. | Self-esteem | 2.44 | 1.09 | 2.37 | 1.00 | | 2.57 | 1.07 | 1.97 | 0.94 | | 2.43 | 1.07 | 0.68** |
| 19. | Self-reproach | 2.91 | 1.02 | 3.02 | 0.94 | | 3.15 | 0.91 | 2.24 | 0.98 | | 2.93 | 1.00 | 0.82 |
| 20. | Self-deprecation | 2.89 | 1.08 | 3.02 | 96.0 | | 3.15 | 0.95 | 2.19 | 1.00 | | 2.92 | 1.05 | 0.84** |
| 30. | Helplessness | 2.69 | 1.12 | 2.64 | 1.06 | | 2.89 | 1.05 | 2.02 | 1.02 | | 2.68 | 1.1 | **08.0 |
| Somatic complaints items | | 20.01 | 4.57 | 19.29 | 3.87 | —1.89 ns | 20.65 | 4.10 | 17.25 | 4.40 | -7.27** | 19.83 | 4.42 | |
| Ξ. | Somatic complaint | 2.28 | 1.03 | 2.19 | 1.00 | | 2.40 | 1.01 | 1.81 | 0.93 | | 2.26 | 1.02 | **89.0 |
| 17. | Pessimism | 2.78 | 1.00 | 2.65 | 0.98 | | 2.86 | 0.98 | 2.40 | 0.97 | | 2.75 | 1.00 | **99:0 |
| 18. | Fatigue | 3.48 | 0.71 | 3.52 | 0.70 | | 3.60 | 0.63 | 3.12 | 0.87 | | 3.49 | 0.72 | .*99:0 |
| 22. | Anger | 2.93 | 0.91 | 2.69 | 0.84 | | 2.97 | 0.88 | 2.57 | 0.91 | | 2.87 | 0.90 | 0.67** |
| 24. | Sleep disturbance | 2.87 | 1.08 | 2.62 | 1.13 | | 2.90 | 1.08 | 2.51 | 1.10 | | 2.81 | 1.10 | **40.0 |
| 27. | Stomachaches | 2.53 | 1.08 | 2.63 | 1.04 | | 2.71 | 1.03 | 2.06 | 1.05 | | 2.56 | 1.07 | 0.68** |
| 28. | Loss of interest | 3.15 | 0.82 | 2.98 | 0.84 | | 3.21 | 0.80 | 2.78 | 0.82 | | 3.11 | 0.83 | 0.59** |
| Fotal score | | 76.51 | 19.25 | 75.65 | 14.81 | $-0.78 \mathrm{ns}$ | 80.65 | 16.46 | 62.58 | 16.87 | 9.44** | 76.30 | 18.26 | |

Table 3. Standardized factor loadings for RADS-2 items, by age-group, sex, and for the whole sample.

| | | 12–17 years (<i>N</i> = 406) | 18–22 years (N = 130) | Females $(N = 407)$ | Males (N = 129) | Whole sample ($N = 536$) |
|---------------------------|-----------------------|-------------------------------|--------------------------|---------------------|-----------------|----------------------------|
| Subscales and items | Abbreviation | , , , | , , | | · · · · · | <u> </u> |
| Dysphoric mood | | | | | | |
| Items | | | | | | |
| 2. | School anxiety | 0.65 | 0.44 | 0.57 | 0.43 | 0.61 |
| 3. | Loneliness | 0.76 | 0.77 | 0.72 | 0.76 | 0.76 |
| 6. | Social withdrawal | 0.75 | 0.81 | 0.71 | 0.80 | 0.75 |
| 7. | Sadness | 0.91 | 0.78 | 0.85 | 0.90 | 0.90 |
| 8. | Crying | 0.80 | 0.54 | 0.63 | 0.86 | 0.76 |
| 16. | Irritability | 0.73 | 0.55 | 0.64 | 0.75 | 0.70 |
| 21. | Self-pity | 0.36 | 0.28 | 0.24 | 0.54 | 0.34 |
| 26. | Worry | 0.75 | 0.54 | 0.62 | 0.73 | 0.71 |
| Anhedonia/negative affect | | | | | | |
| Items | | | | | | |
| 1. | Reduced affect | 0.89 | 0.84 | 0.85 | 0.91 | 0.88 |
| 5. | Self-worth | 0.81 | 0.75 | 0.83 | 0.69 | 0.80 |
| 10. | Anhedonia-peers | 0.64 | 0.56 | 0.59 | 0.64 | 0.62 |
| 12. | Low self-worth | 0.72 | 0.80 | 0.79 | 0.63 | 0.73 |
| 23. | Reduced speech | 0.62 | 0.59 | 0.61 | 0.66 | 0.62 |
| 25. | Anhedonia-general | 0.74 | 0.70 | 0.70 | 0.72 | 0.73 |
| 29. | Appetite disturbance | 0.67 | 0.50 | 0.55 | 0.67 | 0.64 |
| Negative self-evaluation | | | | | | |
| Items | | | | | | |
| 4. | Feelings of rejection | 0.58 | 0.43 | 0.56 | 0.49 | 0.54 |
| 9. | Worthlessness | 0.82 | 0.70 | 0.76 | 0.78 | 0.79 |
| 13. | Discouragement | 0.81 | 0.81 | 0.75 | 0.83 | 0.80 |
| 14. | Self-injurious | 0.81 | 0.75 | 0.75 | 0.79 | 0.79 |
| 15. | Self-esteem | 0.68 | 0.53 | 0.65 | 0.53 | 0.65 |
| 19. | Self-reproach | 0.94 | 0.96 | 0.95 | 0.91 | 0.94 |
| 20. | Self-deprecation | 0.96 | 0.94 | 0.97 | 0.94 | 0.96 |
| 30. | Helplessness | 0.89 | 0.76 | 0.83 | 0.87 | 0.86 |
| Somatic complaints | · | | | | | |
| Items | | | | | | |
| 11. | Somatic complaint | 0.67 | 0.55 | 0.60 | 0.63 | 0.65 |
| 17. | Pessimism | 0.69 | 0.48 | 0.61 | 0.70 | 0.65 |
| 18. | Fatigue | 0.85 | 0.68 | 0.76 | 0.82 | 0.81 |
| 22. | Anger | 0.72 | 0.56 | 0.70 | 0.66 | 0.69 |
| 24. | Sleep disturbance | 0.60 | 0.49 | 0.57 | 0.55 | 0.57 |
| 27. | Stomachaches | 0.67 | 0.54 | 0.61 | 0.53 | 0.64 |
| 28. | Loss of interest | 0.60 | 0.61 | 0.55 | 0.56 | 0.60 |

Note. RADS-2: Reynolds Adolescent Depression Scale second edition.

Table 4. Reliability measures.

| Scale | McDonalds Omega (95% CI |
|------------------------------------|-------------------------|
| Dysphoric mood subscale | 0.84 (0.81 – 0.86) |
| Anhedonia/negative affect subscale | $0.83 \ (0.81 - 0.85)$ |
| Negative self-evaluation subscale | 0.89 (0.87 - 0.90) |
| Somatic complaints subscale | 0.79 (0.75 - 0.81) |
| RADS-2 total scale | 0.85 (0.83 - 0.87) |

Note. RADS-2: Reynolds Adolescent Depression Scale second edition.

scale), all at p < 0.01. Convergent validity was found, with correlations with established measures of depression ranging from 0.84 (MADRS) to 0.92 (BYI-D), all at p < 0.01. Discriminant validity was also established with correlations with measures of distinct constructs ranging from -0.18 (PROMIS-physical activity) to -0.48. (PROMIS-peer-relationships), also all at p < 0.01.

Discussion

In this study, we aimed to test the psychometric properties of the Swedish version of RADS-2, an internationally established measure of depression, in a clinical sample. Acceptable fit indices in the CFA supported the four-factor structure demonstrated in previous studies [11,13,16]. We draw this conclusion despite RMSEA not reaching the

preferred cut-off value, as all other fit indices did. Also, the SRMR is more accurate than RMSEA for model fit evaluation when data is ordinal and when item-level non-normality is identified [54], as was the case in our sample.

The RADS-2 McDonald's Omega ranged from 0.79 to 0.89 which is in line with what has been found previously [10], indicating that the scale is reliable also in this context. We tested four levels of measurement invariance for both sex (males and females), age-group (12-17 years and 18-22 years), and for the clinical/non-clinical samples. The presence of some significant Satorra–Bentler scaled χ^2 difference tests is, strictly interpreted, indicative of measurement non-invariance on that comparison. This is however a sample size-sensitive test and caution is needed to not mistakenly reject null-hypotheses of model equivalence. Across all levels, there were only minimal changes in RMSEA, well within the acceptable boundaries. The changes in CFI were also small; only at the scalar level for clinical-nonclinical samples did the change in CFI touch the recommended threshold (-0.002). Given that this threshold is highly conservative based on having many groups (10-20), whereas in our case we had only two, and considering the delta-RMSEA results; it was considered in general that measurement invariance was demonstrated. RADS-2 is therefore suitable for use in both males and females; in both 12-17- and 18-22-year-olds, and in



Table 5. Measurement invariance goodness-of-fit for the four-factor model of RADS-2, presented with separate CFAs for sex and age-group, as well as invariance models for sex, age-group, and the clinical/non-clinical sample.

| | | | | | | Dit | ference | e tests | |
|---|-----|------------------|-------|----------|--|-----------------|-------------|--------------|----------------|
| | | | Mo | odel fit | | Satorra–Bentler | | | |
| Subgroup CFA | n | $\chi^2(df)$ | CFI | TLI | RSMEA (90% CI) | Satoria Bentier | | | |
| Females | 407 | 1581.09** (399) | 0.926 | 0.919 | 0.085 (0.081 - 0.090) | | | | |
| Males | 129 | 612.40** (399) | 0.953 | 0.949 | 0.065 (0.054 - 0.075) | | | | |
| 12–17 years of age | 406 | 1468.59** (399) | 0.991 | 0.942 | 0.081 (0.077 - 0.086) | | | | |
| 18–22 years of age | 130 | 697.44** (399) | 0.921 | 0.914 | 0.076 (0.067 - 0.085) | | | | |
| Sex-invariance | | $\chi^2(df)$ | CFI | TLI | RSMEA (90% CI) | $\Delta \chi^2$ | Δdf | ΔCFI | Δ RSMEA |
| Configural | | 2027.970** (798) | 0.938 | 0.932 | 0.076 (0.070 - 0.079) | | | | |
| Thresholds constrained | | 2058.407** (828) | 0.938 | 0.935 | 0.075 (0.068 - 0.079) | 28.289 | 30 | 0.000 | -0.001 |
| Thresholds and factor loadings constrained (metric) | | 2043.272** (854) | 0.940 | 0.939 | 0.072 (0.068 - 0.077) | 36.645 | 26 | 0.002 | -0.003 |
| Thresholds, factor loadings, and intercepts constrained (strong/scalar) | | 2008.293** (880) | 0.943 | 0.944 | 0.069 (0.065 – 0.073) | 32.362 | 26 | 0.003 | -0.003 |
| Age group-invariance | | $\chi^2(df)$ | CFI | TLI | RSMEA (90% CI) | $\Delta\chi^2$ | Δdf | ΔCFI | ΔRSMEA |
| Configural | | 1930.718** (798) | 0.956 | 0.952 | 0.073 (0.067 - 0.076) | $\Delta \chi$ | Диг | ДСП | ANJIVILA |
| Thresholds constrained | | 1966.381** (828) | 0.955 | 0.953 | 0.073 (0.067 - 0.076) 0.072 (0.068 - 0.076) | 36.205 | 30 | -0.001 | -0.001 |
| Thresholds and factor loadings constrained (metric) | | 1916.537** (854) | 0.958 | 0.958 | 0.072 (0.003 - 0.070) 0.068 (0.064 - 0.072) | 25.880 | 26 | 0.003 | -0.001 |
| Thresholds, factor loadings, | | 1929.156** (880) | 0.959 | 0.959 | 0.067 (0.063 - 0.071) | 45.257* | 26 | 0.003 | -0.004 |
| and intercepts constrained (strong/scalar) | | 1727.130 (000) | 0.555 | 0.555 | 0.007 (0.003 – 0.071) | 43.237 | 20 | 0.001 | -0.001 |
| Clinical/non-clinical sample invariance ^a | | $\chi^2(df)$ | CFI | TLI | RSMEA (90% CI) | $\Delta \chi^2$ | Δdf | ΔCFI | Δ RSMEA |
| Configural | | 3543.956** (798) | 0.939 | 0.934 | 0.078 (0.069 - 0.081) | ,, | | | |
| Thresholds constrained | | 3610.704** (828) | 0.939 | 0.935 | 0.077 (0.068 - 0.080) | 50.832* | 30 | 0.000 | -0.001 |
| Thresholds and factor loadings constrained (metric) | | 3654.037** (854) | 0.938 | 0.937 | 0.076 (0.068 - 0.079) | 90.572** | 26 | -0.001 | -0.001 |
| Thresholds, factor loadings, and intercepts constrained (strong/scalar) | | 3773.993** (880) | 0.936 | 0.937 | 0.077 (0.074 – 0.079) | 161.201** | 26 | -0.002 | 0.001 |

Note. a Please note that this analysis includes data from a normative sample in addition to the present clinical sample. **p < 0.001; *p < 0.05.

RADS-2: Reynolds Adolescent Depression Scale second edition; χ^2 : Chi-square; df: degrees of freedom; CFI: comparative fit index; TLI: Tucker–Lewis index; RMSEA: root mean square error of approximation; CI: confidence interval; ΔCFI: change in comparative fix index

Table 6. Spearman's correlation coefficients between RADS-2 total scale and RADS-2 subscales, as well as between RADS-2 total scale and validation instruments.

| Measure | Dysphoric mood | Anhedonia/negative affect | Negative self-evaluation | Somatic complaints | RADS-2 total scale |
|-------------------------------------|----------------|---------------------------|--------------------------|--------------------|--------------------|
| RADS-2 Dysphoric mood | | 0.58** | 0.77** | 0.68** | 0.87** |
| RADS-2 Anhedonia/negative affect | | | 0.75** | 0.52** | 0.83** |
| RADS-2 Negative self-evaluation | | | | 0.65** | 0.94** |
| RADS-2 Somatic complaints | | | | | 0.80** |
| BYI-D . | 0.77** | 0.77** | 0.89** | 0.73** | 0.92** |
| MADRS | 0.71** | 0.71** | 0.79** | 0.70** | 0.84** |
| RCADS-depression | 0.72** | 0.70** | 0.80** | 0.75** | 0.86** |
| PROMIS-peer relationships item bank | -0.38** | -0.57** | -0.45** | -0.25** | -0.48** |
| PROMIS-physical activity item bank | -0.14** | -0.18** | -0.15** | -0.14** | -0.18** |

Note. **p < 0.01 (2-tailed).

RADS-2: Revnolds Adolescent Depression Scale second edition; BYI-D: Beck Youth Inventories of Emotional and Social Impairment Depression subscale; MADRS: Montgomery - Asberg Depression Rating Scale; RCADS-depression: Revised Child Anxiety and Depression Scale, depression subscale; PROMIS: Patient Reported Outcome Measurements Information System.

clinical as well as clinical populations, and comparisons of latent mean-scores between these populations are valid.

As expected the RADS-2 scores were higher in the present clinical sample than in the previously published normative sample [20]. Sex-differences were found, with females scoring generally higher than males. We interpret this as a result of the female sample containing more individuals with primary affective diagnoses, reflecting the diagnosis-distribution in the general population [55]. Indeed, the presence of a primary diagnosis of an affective disorder was more common in the females than in the males in our sample, Chi^2 (1 DF, n = 223) = 7.78, p = 0.005, supporting this conclusion. Convergent and discriminative validity were demonstrated with correlations using scales of similar and different constructs.

An unexpected finding was the low correlation between RADS-2 and PROMIS physical activity (-0.18), as associations between depression and physical inactivity have been previously shown [56,57]. In individuals from the present agerange there seems to be poor agreement between selfreported and objectively measured physical activity [58], and self-report bias has been reported particularly in the presence of mental health problems [59].

Another unexpected finding was the low factor loading on item 21 (0.34, see Table 3) rating self-pity: 'I am feeling sorry for myself', a trend that was particularly strong in the females and in the older age-group. This was not seen in our normative sample study [20] and it is therefore unlikely to be culturally or translation-related. It is possible that some individuals with depressive symptoms do not find themselves worthy of love and compassion [60], and not valued enough to feel sorry for. Supporting this is the finding that self-compassion is lower in adolescent girls compared to boys, and lower levels correlate with depression ratings [61].

Limitations and strengths

One limitation of this study is that participation rates were low and therefore the extent to which our sample confidently represents an unselected clinical population is unclear. Most of the participants were recruited from Child and Adolescent Psychiatry, which potentially limits generalizability beyond this context. Also, regarding external validity, only one fourth of the participants were male. This was expected given that internalizing symptoms are less prevalent in young males compared to young females [62], and the sex-distribution in our sample roughly corresponds to that of previous studies [55]. Measurement invariance analysis supported the hypothesis that the scale performs equally well in both sexes, which reduces the impact of the unequal sex distribution in this sample.

Another limitation was that only self-rating was performed. To compare self-rating scores with clinician ratings would have improved the validity-analyses, and clinical assessments would also have increased the reliability of the psychiatric diagnoses that were now extracted from participants' medical records. The low frequency of affective diagnoses in the current sample is a limitation that was likely caused by under-reporting or lack of diagnostic routines. More precise diagnosis-data would have enabled the computation of receiver operating characteristic curves to suggest optimal cutoff-scores for clinical caseness. The limitations of our diagnosis-data as well as the broad inclusion criteria need to be kept in mind when interpreting all results of the study. A potential weakness of RADS-2 is that the scale does not capture all dimensions of depression postulated by the DSM-5, for example attention deficit is not explicitly measured. It is possible that symptoms that are less specific to depression as compared to depressed mood and diminished interest/pleasure [6] have been omitted from the RADS-2 for that reason.

In terms of analysis, validity analyses with simple correlations return only the relationship between variables without quantifying the agreement between the two. More general disadvantages of classical test theory have been elaborated elsewhere [63] and modern item response theory is increasingly being used to evaluate clinical measures in a more sophisticated manner [64]. Therefore, as a future direction we suggest using item-response modeling to evaluate the psychometric properties of RADS-2, even though the dimensionality of the scale would then have to be discarded.

Strengths of this study include the recruitment of patients from a rural to university town area, as well as from different levels of care. Recruiting a mixed clinical sample from an extended age-range including both teenagers and young adults increases the applicability of the scale. This is advantageous given the high frequency of comorbidity [65] and supports future research spanning the age-range from adolescence to young adulthood. As measurement invariance holds for RADS-2 in the whole age-range of this study population, it will be possible to study latent means in this extended population in future studies [47].

Conclusions

RADS-2 displayed good psychometrical properties in the current sample, with supported factor structure, acceptable to good reliability, good validity, and measurement invariance, supporting the view of RADS-2 being a reliable and useful instrument. We conclude that the Swedish version of RADS-2 may be used by clinicians to evaluate symptoms of depression, and by researchers for observational and experimental purposes.

Acknowledgments

The authors would like to thank the study participants, the clinic-staff and the research-assistants who made this study possible.

Disclosure statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

This study was funded by the Swedish Research Council, the County Council of the Region Västerbotten Central ALF under Grant [RV-967045], the County Council of the Region Västernorrland, municipality of Örnsköldsvik and the Kempe foundation under Grant [LVNFOU933598], Lars Jacob Boëthius foundation, the Oskar foundation and the Swedish society of medicine under Grant [SLS-935854]. The funders had no role in the design, methods, subject recruitment, data collection, analysis, or preparation of the manuscript.

ORCID

Erik Ekbäck (h) http://orcid.org/0000-0003-0496-7158 Ida Blomqvist (i) http://orcid.org/0000-0003-0487-5552 Inga Dennhag (b) http://orcid.org/0000-0001-7336-1657 Eva Henje (i) http://orcid.org/0000-0001-5790-0518

Data availability statement

The data that support the findings of this study are available from the corresponding author, EE, upon reasonable request.

References

- Lepine JP, Briley M. The increasing burden of depression. Neuropsychiatr Dis Treat. 2011;7(1):3-7.
- WHO. World health organization: depression and other common mental disorders: global health estimates. Geneva, Switzerland: WHO; 2017.
- Avenevoli S, Swendsen J, He JP, et al. Major depression in the [3] national comorbidity survey-adolescent supplement: prevalence, correlates, and treatment. J Am Acad Child Adolesc Psychiatry. 2015:54(1):37-44.e2.
- Johnson D, Dupuis G, Piche J, et al. Adult mental health out-[4] comes of adolescent depression: a systematic review. Depress Anxiety. 2018;35(8):700-716.
- Leone M, Kuja-Halkola R, Leval A, et al. Association of youth depression with subsequent somatic diseases and premature death. JAMA Psychiatry. 2021;78(3):302.
- APA. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington (VA): American Psychiatric Publishing; 2013.
- Henje Blom E, Ho TC, Connolly CG, et al. The neuroscience and context of adolescent depression. Acta Paediatr. 2016;105(4): 358-365.

- Gogtay N, Giedd JN, Lusk L, et al. Dynamic mapping of human cortical development during childhood through early adulthood. Proc Natl Acad Sci USA. 2004;101(21):8174-8179.
- Ekbäck E, Granåsen G, Svärling R, et al. Clinical effectiveness of training for awareness resilience and action online compared to standard treatment for adolescents and young adults with depression: study protocol and analysis plan for a pragmatic, multi-center randomized controlled superiority trial. Front Psvchiatry. 2021:12:674583-674583.
- [10] Reynolds WM. Reynolds adolescent depression scale. 2nd ed. Professional manual. Odessa (FL): Psychological Assessment Resources; 2002.
- [11] Fonseca-Pedrero E, Wells C, Paino M, et al. Measurement invariance of the Reynolds depression adolescent scale across gender and age. Int J Test. 2010;10(2):133-148.
- [12] 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. 2016;7:208.
- [13] 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(3):642-651.
- [14] King CA, O'Mara RM, Hayward CN, et al. Adolescent suicide risk screening in the emergency department. Acad Emerg Med. 2009; 16(11):1234-1241.
- [15] 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.
- [16] Walker L, Merry S, Watson PD, et al. The Reynolds adolescent depression scale in New Zealand adolescents. Aust N Z J Psychiatry. 2005;39(3):136-140.
- [17] Osman A, Gutierrez PM, Bagge CL, et al. Reynolds adolescent depression scale-second edition: a reliable and useful instrument. J Clin Psychol. 2010;66(12):1324-1345.
- Sami S, Ahmad R, Khanam SJ. Translation of Reynolds adolescent [18] depression scale - second edition in Pakistan: reliability estimates and factor analysis. Pak J Clin Psychol. 2013;12(2):19-32.
- [19] Carillo A. Creation and use of the RADS-2TM and the RCDSTM-2:SF: Spanish translations (white paper). PAR-Inc. Lutz, Florida. 2020.
- [20] Blomqvist I, Ekbäck E, Dennhag I, et al. Validation of the Swedish version of the Reynolds Adolescent Depression Scale second edition (RADS-2) in a normative sample. Nord J Psychiatry. 2020;75: 292-300
- [21] Beck JS, Beck AT, Jolly J. Manual for the beck youth inventories of emotional and social impairment. San Antonio (TX): The Psychological Corporation; 2001.
- [22] Bose-Deakins JE, Floyd RG. A review of the beck youth inventories of emotional and social impairment. J School Psychol. 2004; 42(4):333-340.
- [23] Steer RA, Kumar G, Beck JS, et al. Evidence for the construct validities of the Beck Youth Inventories with child psychiatric outpatients. Psychol Rep. 2001;89(3):559-565.
- Dunerfeldt ME, Söderström A. Bengt. Bedömningsinstrument [24] inom BUP i Stockholm. Kartläggning och faktasammanställning. Stockholm: Barn- och ungdomspsykiatri, Stockholms läns landstina: 2010.
- [25] SBU. Diagnostik och uppföljning av förstämningssyndrom. En systematisk litteraturöversikt. SBU-rapport nr 212. Stockholm: SBU, Statens beredning för medicinsk utvärdering. 2012.
- [26] Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry. 1979;134:382-389.
- [27] Carmody TJ, Rush AJ, Bernstein I, et al. The Montgomery Asberg and the Hamilton ratings of depression: a comparison of measures. Eur Neuropsychopharmacol. 2006;16(8):601-611.
- [28] Williams JB, Kobak KA. Development and reliability of a structured interview guide for the Montgomery Asberg Depression Rating Scale (SIGMA). Br J Psychiatry. 2008;192(1):52-58.

- Ntini I, Vadlin S, Olofsdotter S, et al. The Montgomery and Åsberg Depression Rating Scale - self-assessment for use in adolescents: an evaluation of psychometric and diagnostic accuracy. Nord J Psychiatry. 2020;74(6):415-422.
- [30] Chorpita BF, Yim L, Moffitt C, et al. Assessment of symptoms of DSM-IV anxiety and depression in children: a revised child anxiety and depression scale. Behav Res Ther. 2000;38(8):835-855.
- [31] Piqueras JA, Martín-Vivar M, Sandin B, et al. The revised child anxiety and depression scale: a systematic review and reliability generalization meta-analysis. J Affect Disord. 2017;218:153-169.
- [32] Giannopoulou I, Pasalari E, Bali P, et al. Psychometric properties of the revised child anxiety and depression scale in Greek adolescents. Clin Child Psychol Psychiatry. 2022;27(2):424-438.
- [33] Kösters MP, Chinapaw MJ, Zwaanswijk M, et al. Structure, reliability, and validity of the revised child anxiety and depression scale (RCADS) in a multi-ethnic urban sample of Dutch children. BMC Psychiatry. 2015;15:132.
- [34] Wolpert M, Cheng H, Deighton J. Measurement issues: review of four patient reported outcome measures: SDQ, RCADS, C/ORS and GBO - their strengths and limitations for clinical use and service evaluation. Child Adolesc Ment Health. 2015;20(1):63-70.
- Patient-reported outcomes in performance measurement. Research Triangle Park, North Carolina, USA: RTI Press/RTI international. 2015.
- PROMIS-Organization. PROMIS pediatric banks item definitions. [36] Unpublished work. 2018. Available from: www.promishealth.org
- Blomqvist I, Chaplin JE, Nilsson E, et al. Swedish translation and cross-cultural adaptation of eight pediatric item banks from the patient-reported outcomes measurement information system (PROMIS)(®). J Patient Rep Outcomes. 2021;5(1):80.
- [38] Carlberg Rindestig F, Wiberg M, Chaplin JE, et al. Psychometrics of three Swedish physical pediatric item banks from the Patient-Reported outcomes measurement information system (PROMIS)®: pain interference, fatigue, and physical activity. J Patient Rep Outcomes. 2021;5(1):105.
- Alavi M, Visentin DC, Thapa DK, et al. Chi-square for model fit in [39] confirmatory factor analysis. J Adv Nurs. 2020;76(9):2209-2211.
- [40] Hooper D, Coughlan J, Mullen MR. Structural equation modelling: guidelines for determining model fit. Electron J Business Res Methods. 2008;6(1):53-60.
- [41] Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Eq Model. 1999;6(1):1-55.
- [42] Li CH. The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. Psychol Methods. 2016;21(3):369-387.
- [43] Li CH. Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. Behav Res. 2016;48(3):936-949.
- [44] Dunn TJ, Baguley T, Brunsden V. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. Br J Psychol. 2014;105(3):399-412.
- Nunally J, Bernstein IH. Psyhometric theory. 3rd ed. New York (NY): McGraw-Hill; 1994.
- Wu H, Estabrook R. Identification of confirmatory factor analysis [46] models of different levels of invariance for ordered categorical outcomes. Psychometrika. 2016;81(4):1014-1045.
- [47] 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 Eq Model. 2020;27(1):111-130.
- [48] Satorra A, Bentler PM. Ensuring positiveness of the scaled difference chi-square test statistic. Psychometrika. 2010;75(2):243-248.
- Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates, Publishers; 1988, p. 567.
- R-team. R: a language and environment for statistical computing. [50] Vienna, Austria: R Foundation for Statistical Computing; 2019.
- [51] Yves R. Lavaan: an R package for structural equation modeling. J Stat Softw. 2012;48(2):1-36.

- [52] SCB. Statistics Sweden (SCB) the Swedish socio-economic classification (SEI). 1984. Available from: https://www.scb. se/contentassets/22544e89c6f34ce7ac2e6fefbda407ef/english ov9999 1982a01 br x11op8204-3.pdf
- SCB. Statistics Sweden (SCB) SEI, yrkesförteckning version 2020-04-2020. Available from: https://www.scb.se/contentassets/ 22544e89c6f34ce7ac2e6fefbda407ef/sei_index_webb_20200416.pdf
- Shi D, Maydeu-Olivares A, Rosseel Y. Assessing fit in ordinal factor analysis models: SRMR vs. RMSEA. Struct Eq Model. 2020;
- [55] Socialstyrelsen. Utvecklingen av psykisk ohälsa bland barn och unga vuxna: Till och med 2016. 2017. Available from: https:// www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/statistik/2017-12-29.pdf
- [56] McMahon EM, Corcoran P, O'Regan G, et al. Physical activity in European adolescents and associations with anxiety, depression and well-being. Eur Child Adolesc Psychiatry. 2017;26(1):111-122.
- [57] Carter T, Morres ID, Meade O, et al. The effect of exercise on depressive symptoms in adolescents: a systematic review and metaanalysis. J Am Acad Child Adolesc Psychiatry. 2016;55(7):580-590.
- [58] Pinto AJ, Roschel H, Benatti FB, et al. Poor agreement of objectively measured and self-reported physical activity in juvenile

- dermatomyositis and juvenile systemic lupus erythematosus. Clin Rheumatol. 2016;35(6):1507-1514.
- [59] Curtis RG, Olds T, Plotnikoff R, et al. Validity and bias on the online active Australia survey: activity level and participant factors associated with self-report bias. BMC Med Res Methodol. 2020; 20(1):6.
- [60] Marsh IC, Chan SWY, MacBeth A. Self-compassion and psychological distress in adolescents-a meta-analysis. Mindfulness (NY). 2018;9(4):1011-1027.
- 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 Psychiatr Psychol. 2020;8: 70-80.
- [62] Blomqvist I, Henje Blom E, Hagglof B, et al. Increase of internalized mental health symptoms among adolescents during the last three decades. Eur J Public Health. 2019;29:925-931.
- [63] Theresa K. Classical test theory: assumptions. In: Equations, limitations, and item analyses. Thousand Oaks (CA): SAGE Publications, Inc; 2005. p. 91.
- Thomas ML. Advances in applications of item response theory to clinical assessment. Psychol Assess. 2019;31(12):1442-1455.
- Thapar A, Collishaw S, Pine DS, et al. Depression in adolescence. [65] Lancet. 2012;379(9820):1056-1067.