The Screen for Child Anxiety Related Emotional Disorders (SCARED): Informant discrepancy, measurement invariance, and test-retest reliability

Brigid Behrens[1],[*], **Caroline Swetlitz**[1],[*], **Daniel S. Pine**[1], **David Pagliaccio**[2]

[1]Emotion and Development Branch, The National Institute of Mental Health, Bethesda, MD 20892

[2]Division of Child and Adolescent Psychiatry, New York State Psychiatric Institute, New York, NY 10032

Abstract

The Screen for Child Anxiety Related Emotional Disorders (SCARED) is a measure widely used to assess childhood anxiety based on parent and child report. However, while the SCARED is a reliable, valid, and sensitive measure to screen for pediatric anxiety disorders, informant discrepancy can pose clinical and research challenges. The present study assesses informant discrepancy, measurement invariance, test-retest reliability, and external validity of the SCARED in 1,092 anxious and healthy parent-child dyads. Our findings indicate that discrepancy does not vary systematically by the various clinical, demographic, and familial variables examined. There was support for strict measurement invariance, strong test-retest reliability, and adequate external validity with a clinician-rated measure of anxiety. These findings further support the utility of the SCARED in clinical and research settings, but low parent-child agreement highlights the need for further investigation of factors contributing to SCARED informant discrepancy.

Keywords

SCARED; psychometrics; anxiety; informant discrepancy; reliability

Introduction

It is considered best practice for clinicians and researchers to utilize information from multiple informants when assessing symptoms of pediatric psychopathology [1]. While this generates a more comprehensive understanding of symptomology, systematic differences occur in the ways that respondents report symptoms [2]. Issues related to discrepant

Correspondence: Brigid Behrens, National Institute of Mental Health, Building 10, Room B1D43, 9000 Rockville Pike, Bethesda, MD 20892, brigid.behrens@nih.gov, Telephone: (301) 402-2352, Fax: (301) 480-4684.
[*]These authors contributed equally to the work.

reporting have begun to be explored for the Screen for Child Anxiety and Related Emotional Disorders - Parent and Child versions, (SCARED-P/C) [3, 4] a dual informant, gold-standard measure of pediatric anxiety symptoms. While the SCARED has been established as a valid, reliable, and sensitive measure of anxiety, prior studies have yielded inconsistent findings regarding informant (parent-child) agreement/discrepancy, with estimates of agreement ranging from $r \sim .2 - .6$ on both the SCARED total score and its subscales [3, 5, 6]. The present study assesses informant discrepancy on the SCARED in the largest sample to date and probes potential clinical, demographic, and familial correlates of discrepant reporting as well as potential psychometric contributors to informant discrepancy.

Previous literature posits that informant discrepancy regarding pediatric anxiety may vary systematically with a variety of demographic and clinical variables. A child's age is one of the most consistent predictors of discrepant reporting. Reports from younger children are more discrepant from parent- or teacher-reports compared to those of older children [6, 7, 8, 9, 10]. Research investigating informant discrepancy as a function of child's sex have yielded mixed results. Several studies have found no significant differences in informant agreement between girls and boys [6, 8, 11, 12], while others report moderate differences [7, 10, 13, 14]. Ethnicity [8], socioeconomic status (SES [8, 14]), parental psychopathology, and a variety of other family-related factors [7, 9, 14] have also been investigated as potential predictors of discrepancy, though this research has been sparser. Recently, Rappaport and colleagues explored informant discrepancy and discriminant validity on the SCARED [5]. While psychiatrically healthy children over-reported symptoms compared to their parents, children with a clinically diagnosed anxiety disorder under- or equally-reported symptoms relative to their parent. Further, informant discrepancy was significantly larger among youth with social anxiety disorder (SAD) relative to those with generalized anxiety disorder (GAD) or comorbid GAD and SAD. Further work is needed to characterize patterns of informant discrepancy and specific demographic, clinical, and psychometric factors that relate to discrepant reporting on the SCARED.

Measurement invariance and test-retest reliability are two such psychometric factors that may contribute to informant discrepancy. Tests of measurement invariance assess the extent to which different groups (e.g., informants or diagnostic groups) interpret questionnaire items similarly [15]. Establishing strict measurement invariance across informants and groups, a critical step in examining a questionnaire's interpretability, suggests that the same underlying constructs contribute to the interpretation of items across groups [16]. The only prior study examining measurement invariance on the SCARED across parent-child dyads ($N = 408$; children were seeking treatment at an outpatient mental health facility) tested four standard levels of measurement invariance and found evidence for partial threshold invariance. This is an important step that warrants replication in a large sample including both patients with anxiety disorders and healthy volunteers [8]. Establishing strict measurement invariance across informants on the SCARED is beneficial for the continued use of the questionnaire.

In addition, issues with questionnaire reliability could potentially contribute to informant discrepancy. In their relatively small initial validation study ($N = 88$ children, $N = 86$ parents), Birmaher and colleagues [3] established strong test-retest reliability of the

SCARED over a 5-day to 15-week window (intraclass correlation coefficient [ICC] = .70 – .90). Since then, Boyd and colleagues [17] found moderate test-retest reliability ($r = .47$) of SCARED child-report over a 6-month period. Additionally, studies examining translated versions of the revised 41-item SCARED have found mostly strong test-retest reliability, though generally over a much shorter time window of 7 – 14 days [18, 19, 20]. However, given the use of the SCARED in studies of the treatment of pediatric anxiety, it is important to establish test-retest reliability of the revised SCARED questionnaire over a longer window of time closer to the length of standard treatment. Additionally, no study to our knowledge has assessed the stability of informant discrepancy on the SCARED over time.

Building on recommendations and limitations from previous work, the aims of the current study were to assess informant agreement/discrepancy on the SCARED in the largest sample to date ($N = 1,092$ parent-child dyads) and to elucidate clinical, demographic, familial, and psychometric factors contributing to informant discrepancy. Specifically, we quantified informant agreement and examined associations between four metrics of informant discrepancy and demographic factors as well as differences as a function of diagnosis (healthy vs. anxious youth). Further, we examined measurement invariance of the SCARED across informants and across ages for parent- and child-report in a sample missing no item-level data. Additionally, we investigated test-retest reliability of SCARED scores and of informant discrepancy. Finally, we examined associations with clinician-rated anxiety to explore external validity.

Methods

Sample and Setting

Participants included 1,092 youth (7–18 years old) and parent dyads who completed the SCARED questionnaire. Individuals were enrolled in IRB-approved research protocols examining pediatric anxiety disorders at the National Institute of Mental Health (NIMH). Child participants and their parents provided written assent and consent, respectively. All participants were assessed using a structured diagnostic interview (Kiddie Schedule for Affective Disorders and Schizophrenia- Present and Lifetime version; K-SADS [21]). Of this sample, 457 youths were seeking treatment for an anxiety disorder and received a primary diagnosis of either generalized anxiety disorder, separation anxiety disorder, or social anxiety disorder. Healthy youth had no current or past psychiatric diagnoses. All child participants had an IQ > 70 and were medication-free; the presence of current major depressive disorder, obsessive compulsive disorder, or post-traumatic stress disorder were exclusionary. Dyads were selected for the current analyses if both parent and child completed the SCARED: with no missing item responses, within two months of one another, and prior to the child beginning treatment at the NIMH.

A subset of these participants was included in a test-retest analysis if they completed a second administration of the SCARED (5 days to 15 weeks after the first administration), prior to the start of treatment with no more than four items missing on the second administration. This subsample included 339 parent-report forms and 359 child-report forms ($n = 298$ complete dyads). The timeframe mirrored Birmaher and colleagues' initial SCARED reliability study [3].

Measures

Child anxiety symptoms were assessed using the SCARED parent and child versions. The SCARED-P and SCARED-C each consist of 41 items that assess a child's recent anxiety symptoms. Participants respond on a 3-point Likert scale of 0 (Not True or Hardly Ever True), 1 (Somewhat or Sometimes True), or 2 (Very True or Often True). Prior confirmatory factory analyses suggest that the instrument measures five distinct domains of anxiety [4, 8, 22]. Thus, in addition to total scores, five subscales were examined: generalized anxiety symptoms (nine items), separation anxiety symptoms (five items), social anxiety symptoms (eight items), panic or somatic symptoms (seven items), and school avoidance (three items). A total score of 25 or above has been suggested to indicate the presence of clinically significant anxiety [3, 5].

Clinicians rated children's anxiety severity during the previous week using the Pediatric Anxiety Rating Scale (PARS), a 50-item checklist examining symptoms of social, separation, and generalized anxiety, specific phobias, and physical symptoms [23]. Recent studies have found the PARS to be psychometrically reliable and valid, and it has been used as an outcome measure for several treatment studies [24–26]. Clinicians integrated parent and child report during interview assessment to rate seven areas of anxiety severity (number of symptoms, frequency, severity of distress associated with anxiety symptoms, severity of physical symptoms, avoidance, interference at home, and interference outside of home). A score of 3 on each of these 5-point scales reflects a clinically significant level of anxiety. Composite PARS scores were calculated by summing 5 of the 7 items (number of symptoms and severity of physical symptoms were excluded as they are likely less related to overall anxiety severity and tend to be highly skewed). PARS scores were only available for a subset of ($n = 213$) youth with a diagnosed anxiety disorder, and all scores reflected anxiety prior to beginning treatment at NIMH.

Children's age, sex, ethnicity, and family socioeconomic status (SES) were assessed using a demographics questionnaire. Highest parental educational attainment and annual income were used as markers of SES. The Weschler Abbreviated Scale of Intelligence II (WASI [27]) was used to assess child IQ. The Family Risk Factor Checklist (FRFC [28]) is a 48-item measure that assesses children's exposure to environmental/family-related risk (five subscales: adverse life events & instability, family structures & SES, parenting practices, parental verbal conflict, and mood problems). Higher scores indicate greater exposure to family risk factors. Demographic and clinical characteristics of the sample are summarized in Table 1.

Analysis

Informant Agreement/Discrepancy: Parent-child agreement was indexed using intra-class correlation (ICC) for SCARED total scores and for each subscale. ICC1 values were calculated using the *psych* package [35] in R v3.3.1 [36] for the whole sample and separately for the healthy and anxious subsamples (Table 2). Four measures were used to characterize informant discrepancy. Raw discrepancy scores (RDS; *raw parent – raw child score*) characterize the magnitude and direction of discrepancy. Absolute values of these RDS (absRDS) characterize the overall magnitude of discrepancy regardless of

directionality. Given that raw discrepancy scores tend to correlate with overall level of symptoms and following recommendations by De Los Reyes and Kazdin, [2] standardized mean difference scores (SMDS) were also calculated by separately z-scoring parent and child scores and then creating a within-dyad difference score (*Z parent – Z child scores).* This approach aids in interpretation by centering and normalizing variance for each informant group. Additionally, absolute values of these SMDS (absSMDS) were calculated to assess the overall magnitude of discrepancy regardless of directionality. Differences between RDS and SMDS tend to be most salient when variance differs by informant groups. As variance was largely similar across informants in this sample, these measures yielded largely convergent results. Nonetheless, these different measures were presented to address concerns about the advantages and disadvantages of RDS and SMDS as well as to improve interpretability and comparison to other studies. A summary of all four measures of discrepancy (RDS, absRDS, SMDS, absSMDS) across the sample and by diagnostic group is presented in Table 1.

Associations between informant discrepancy and demographic variables of interest were assessed using independent samples t-tests for binary predictors (sex, ethnicity) and Spearman's rho correlations for continuous predictors (age, SES, IQ, FRFC) to account for potential deviations from normality among predictors and the ordinal nature of the SES variables. Independent samples t-tests were also used to compare RDS, absRDS, SMDS, and absSMDS between healthy and anxious parent-child dyads.

Measurement Invariance: Next, we conducted three tests of measurement invariance using confirmatory factor analysis (CFA) in MPlus [29]. The first analysis built on prior work [8] and used a within-subjects model to assess invariance across parent- and child-report. Further, to assess potential effects of age, we conducted separate multi-group measurement invariance models to test invariance of the SCARED-C, comparing younger and older children, and invariance of the SCARED-P, comparing parents of younger and older children. Specifically, dyads were separated into two groups at the median sample age of 12 ($n = 557 < 12$ years old and $n = 535$ 12 years old).

All CFAs used a mean- and variance-adjusted weight least squares (WLSMV) estimator to account for the ordinal nature of the SCARED response scale [8, 30]. Additionally, as in the prior study of invariance on the SCARED [8], we fixed factor means to zero, factor variances to one, and residual variances to one to ensure model identification for both parents and youths [31]. Each measurement invariance test included CFAs at four levels of increasing stringency. First, we examined configural invariance, which tests the factor structure across the two groups/informants. The second level tested weak or metric invariance, locking factor loadings to be equal across groups/informants. The third level tested strong or threshold invariance, locking factor loadings and item thresholds to be equal across groups/informants. The fourth level tested strict or residual invariance, locking item loading, thresholds, and residual item invariance to be equal across groups/informants. Given the oversensitivity of χ^2 tests for model fit and model comparison (e.g. Mplus difference test function) to large sample sizes as in this study, we report χ^2 values but rely on other measures of model fit. Specifically, good model fit was established by a comparative fit index (CFI) > .95 and root mean square error of approximation (RMSEA)

< .06. Measurement invariance at each level was established by small changes in model fit, specifically a decrement in CFI < .01 and an increase in RMSEA < .015 [32, 33, 34].

Test-Retest: We examined the test-retest reliability of SCARED scores and of informant discrepancy (SMDS; recalculated in this subsample). Reliability was assessed using linear mixed effects models using the *lme4* package [37] in R. These models included within-subjects effects for participant and timepoint, controlling for the number of days between assessments. ICC values were extracted, indicating the proportion of participant-specific to total variance. Twelve test-retest reliability models were tested examining total and five subscale scores from parent report ($n = 339$) and from child report ($n = 359$). Test-retest reliability of informant discrepancy (SMDS of total and subscale scores) was assessed in $n = 298$ dyads that completed the SCARED at two timepoints.

External Validity: Finally, we examined the associations between parent/child report on the SCARED and clinician-rated PARS scores. First, we assessed Pearson's correlations between the PARS, the SCARED-P, SCARED-C, and the SCARED-P/C mean scores. Next, we conducted a multiple linear regression to assess whether SCARED-P and SCARED-C predicted unique or shared variance in PARS scores.

Results

Informant Agreement/Discrepancy:

Weak but significant informant agreement was found in the full sample between parent- and child-report on the SCARED (total and subscale scores ICCs = .14 −.19; Table 2). In the anxious subsample, the SCARED-P and SCARED-C showed weak but significant agreement for total scores and for all subscales besides the Panic subscale. Within the healthy subsample, parent-child agreement was only significant for the Social Phobia subscale.

An independent samples t-test indicated significant differences in informant discrepancy between healthy and anxious participants (Table 1) on all four measures: raw difference scores (RDS), absolute values of raw difference scores (absRDS), standardized mean difference scores (SMDS), and absolute values of standardized mean difference scores (absSMDS). While parents under-reported symptoms relative to their child across both groups, the magnitude of the discrepancy was significantly greater in dyads with an anxious child. It is important to note, however, that this could be due to the relatively smaller range of scores present in the non-anxious group.

There were few meaningful correlations between the demographic variables of interest and informant discrepancy (Table 3). Age showed a weak negative correlation with absSMDS. IQ and income had weak, but statistically significant positive associations with both SMDS and RDS. Sex differences in SMDS were also identified such that females (SMDS: $M = -.08$, $SD = 1.25$) reported more symptoms relative to their parents whereas males (SMDS: $M = .08$, $SD = 1.30$) reported less symptoms than their parents. Of the variables investigated, parental education was the strongest predictor of discrepancy (RDS and SMDS), with more highly educated parents reporting more symptoms than their children. Ethnicity (white vs.

non-white) and differences in familial risk factors were not significant predictors of discrepant reporting.

Next, we conducted exploratory regression analyses (Table 4) to examine how much variance in discrepancy scores was accounted for jointly by all factors (excluding the Family Risk Factor Checklist (FRFC) to maintain a larger analysis sample size of $n = 730$ dyads). Age, sex, ethnicity, highest parental education, income, child IQ, SCARED-P/C mean scores, and diagnosis accounted for 8.27% of the variance in RDS ($R^2 = .08$, $F(8, 721) = 8.12$, $p < .001$) and 7.41% of the variance in SMDS ($R^2 = .07$, $F(8, 721) = 7.22$, $p < .001$). These factors accounted for 17.64% of the variance in absRDS ($R^2 = .18$, $F(8, 721) = 19.31$, $p < .001$) and 19.26% of the variance in absSMDS ($R^2 = .19$, $F(8, 721) = 21.50$, $p < .001$). Most of this was accounted for by positive associations with SCARED-P/C mean scores in all four regressions. RDS and SMDS differed by child's diagnosis. Parental education predicted RDS, absRDS, and SMDS.

Measurement Invariance:

Examining the five-factor configural model across informants indicated good model fit (Table 5). Strict measurement invariance was found between parent and child SCARED reports, as evidenced by below threshold changes in CFI and RMSEA at each level of invariance. Next, in two separate models, strict invariance was found for parent- and for child-report splitting the sample at the median child age of 12 years old (Table 6).

Test-Retest Reliability:

From the available sample, $n = 339$ parents and $n = 359$ children completed a second administration of the SCARED that fit the criteria noted in the methods. Parent-report forms were completed an average of 38.62 days apart ($SD = 22.24$) and child-report forms were completed an average of 40.29 days apart ($SD = 23.72$). SCARED total and subscale scores showed moderate to excellent test-retest reliability (Table 7). Specifically, children showed acceptable reliability across total and subscale scores (ICC = .59 – .61), while parent-report showed higher reliability (ICC = .74 – .86). Informant discrepancy (SMDS) was also moderately reliable over time (ICC = .59 – .66; Table 7).

External Validity:

A subset of $n = 201$ anxious patients were assessed by clinician interview using the Pediatric Anxiety Rating Scale (PARS; $M = 15.8$, $SD = 4.09$). SCARED-P ($r = .25$, $p < .001$), SCARED-C ($r = .23$, $p < .001$), and the SCARED-P/C mean ($r = .32$, $p < .001$) scores correlated significantly with PARS scores. In a multiple regression analysis, SCARED-C and SCARED-P explain 10% of the variance in PARS score ($R^2 = .10$, $F(2,198) = 11.54$, $p < .001$). Both parent- ($\beta = .22$, $t = 3.39$, $p < .001$) and child-report ($\beta = .20$, $t = 3.00$, $p < .005$) on the SCARED significantly predicted unique variance in PARS scores. As expected, collinearity between SCARED-C and SCARED-P was low (variance inflation factor = 1.01).

Discussion

Informant Discrepancy:

The current findings indicate that the association between parent and child reports of anxiety on the SCARED questionnaire is on the lower end of prior estimates [5, 6, 10, 38]. Parents tend to report less symptoms than their children overall and this discrepancy was more pronounced in parent-child dyads with a clinically anxious child. However, it should be noted that this finding could be due in part to a "floor effect" in the healthy volunteer group, where anxiety severity was generally low. Despite marked informant discrepancy, this did not appear to vary systematically based on the demographic variables of interest. All significant correlations between predictors and different measures of informant discrepancy were weak, and the large sample size contributed to statistical significance. These results differ from previous findings suggesting that myriad demographic and family-related variables more strongly predict discrepant reporting on the SCARED [6, 8, 9, 12]. The findings of this study support the utility of combining parent and child SCARED scores to obtain a comprehensive view of anxiety symptomology without introducing systematic bias from extraneous factors. Nonetheless, given low informant agreement, more research is warranted into what other factors may contribute to discrepancy, particularly given the reliability of discrepancy scores over time.

Measurement Invariance:

The current results support strict measurement invariance across informants in a large sample missing no item-wise SCARED data. These results are supported by below threshold changes in CFI and RMSEA across the four levels of invariance. While the χ^2 difference tests were significant, these tend to be inflated with large sample sizes and thus we rely on CFI and RMSEA as measures of model fit. These findings largely support and expand on previous literature. Dirks and colleagues established partial invariance (freeing 22 item thresholds) between parent and child informants in a large, but considerably smaller sample ($N = 408$) [8]. Based on the current data and Dirks and colleagues' prior findings, we argue that the SCARED likely exhibited strict invariance across informant, i.e. parent- vs. child-report.

We also found evidence for strict measurement invariance between younger and older children and between the parents of younger and older children. In examining invariance of child-report across younger/older children, changes in CFI and RMSEA were below the set thresholds. However, comparing report from parents of younger/older children, the change in CFI from the strong to strict invariance model marginally exceeded the set threshold of .01. That said, the change in RMSEA was below the established threshold. These data suggest that the interpretation of the SCARED is not significantly impacted by the age of the child. This work may be further expanded in the future using newer approaches allowing for testing invariance along age as a continuous covariate, e.g. Bauer (2017) [39]. Currently, this is not implemented for models fit with a mean- and variance-adjusted weight least squares (WLSMV) estimator, which we use given the ordinal nature of the SCARED items and to maintain consistency with prior work.

Test-Retest Reliability:

Since the initial SCARED reliability study [3], there has been limited research on the test-retest reliability of the revised SCARED questionnaire [4], and no previous studies have explored the reliability of the parent-child discrepancy over time. Reliability of the parent-report was higher than for child-report; however, both showed moderate to high ICC values, based on prior guidelines [40, 41]. This suggests that individuals respond similarly over time, which further supports the use of the SCARED as a stable measure of anxiety. Interestingly, informant discrepancy, i.e. the amount that informants disagree, also remained moderately consistent over time.

External Validity:

Of note, in a subset of anxious patients, we found correlations between SCARED-P and SCARED-C scores and the Pediatric Anxiety Rating Scale (PARS) that were lower than prior estimates (e.g. $r > .32$ [23]). This could be related to methodological differences between the two studies. The prior study [23] completed all measures on one day, whereas measures could be completed on different days in the current study. Regardless, both SCARED-P and SCARED-C did predict unique variance in clinician-rated anxiety severity on the PARS. This suggests that the SCARED-P and the SCARED-C may capture some meaningfully different aspects of the child's anxiety symptoms.

Despite the presence of statistically significant correlations, the low magnitude of these correlations is a concern that could be addressed through additional research on informant discrepancy in the assessment of anxiety. Future research could examine factors that influence associations among the SCARED-P, SCARED-C, and clinician-rated anxiety as well as explore associations with biological measures and other clinical measures, such as long-term outcome. Alternatively, novel assessment techniques that harness digital technology could be explored. Continued research in these and other areas may clarify the meaning of informant discrepancy.

There were several key limitations to our study. First, this was a secondary data analysis, limiting our analyses to the existing data. This presented us with several distinct issues. First, we did not have data on which parent completed the questionnaire. As such, we were unable to examine whether mothers or fathers were more or less discrepant with child-report. Recent work by Jansen et al. [42] suggests that mothers are less discrepant with their child relative to fathers, illustrating the need for further research. Another shortcoming was that not all participants completed every form. Because only a smaller subset of our large sample also completed the FRFC ($n = 358$) and PARS ($n = 213$), we suggest that future research continue to examine these variables' associations with informant discrepancy on the SCARED.

Recent work also suggests that the SCARED factor structure may differ as a function of informant ethnicity, although these findings are inconsistent [8, 17, 43, 44]. These conflicting findings could reflect cultural differences in the presentation, understanding, and stigma associated with these symptoms. Unfortunately, due to the largely homogenous nature of our sample, we were unable to explore invariance as it relates to ethnicity. In

addition, the stringent exclusionary criteria (i.e. the requirement that all participants be medication free, present with no co-occurring major depressive disorder, obsessive compulsive disorder, or post-traumatic stress disorder, have an IQ > 70, and present prior to treatment) influences generalizability of these findings. Future research should continue to examine informant discrepancy, measurement invariance, and test-retest reliability of the SCARED in diverse samples to expand replicability and generalizability of findings.

In sum, using the largest sample to date, our clinical, demographic, and psychometric findings further support the reliability and validity of the SCARED. While measurement invariance analyses suggest that parents and children use and interpret the scale in similar ways, it is noteworthy that lower levels of informant agreement were observed in our sample compared to previous studies. These findings hold important clinical significance, supporting the use of the SCARED as a psychometrically valid tool for self- and parent-report of anxiety symptoms in children, but also highlight the need for further study of the determinants of informant discrepancy and the unique information captured by the parent- and child-report on the SCARED relative to clinician interview.

Summary

Self-report measures are a critical tool in psychological and psychiatric research that can offer insight into an individual's level and characteristics of impairment. The Screen for Child Anxiety Related Emotional Disorders (SCARED) is one of the most commonly used questionnaires for assessing childhood anxiety. While the SCARED is a reliable, valid, and sensitive measure to screen for pediatric anxiety disorders, informant discrepancy can pose clinical and research challenges [2]. In a sample of $N = 1,092$ anxious and healthy parent-child dyads, variables such as child's age, sex, socioeconomic status, symptom severity, and family stress did not systematically predict discrepant reporting. Further, the SCARED showed strict measurement invariance, strong test-retest reliability, and the SCARED-C and SCARED-P predicting unique variance in a clinician-rated measurement of anxiety. These findings suggest not only that item interpretation is not responsible for rater discrepancy, but also that the SCARED-C and SCARED-P may capture meaningful differences in the child's anxiety symptoms. Given the widespread use of the SCARED by both practitioners and researchers, an understanding of its psychometric properties and potential factors driving discrepant reporting is important.

Funding:

References

1. Silverman WK, Ollendick TH (2008) Child and adolescent anxiety disorders In: Hunsley J, Mash EJ (eds) A guide to assessments that work. Oxford Univ Press, NY, pp 181–206

2. De Los Reyes A, Kazdin AE (2004) Measuring informant discrepancies in clinical child research. Psychol Assess 16: 330–334 [PubMed: 15456389]

3. Birmaher B, Khetarpal S, Brent D, Cully M, Balach L, Kaufman J et al. (1997) The screen for child anxiety related emotional disorders (SCARED): scale construction and psychometric characteristics. J Am Acad Child Adolesc Psychiatry 36: 545–553 [PubMed: 9100430]

4. Birmaher B, Brent DA, Chiappetta L, Bridge J, Monga S, Baugher M (1999) Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): a replication study. J Am Acad Child Adolesc Psychiatry 38: 1230–1236 [PubMed: 10517055]

5. Rappaport BI, Pagliaccio D, Pine DS, Klein DN, Jarcho JM (2017) Discriminant validity, diagnostic utility, and parent-child agreement on the Screen for Child Anxiety Related Emotional Disorders (SCARED) in treatment-and non-treatment-seeking youth. J Anxiety Disord 51: 22–31 [PubMed: 28886420]

6. Becker EM, Jensen-Doss A, Kendall PC, Birmaher B, Ginsburg GS (2016) All anxiety is not created equal: correlates of parent/youth agreement vary across subtypes of anxiety. J Psychopathol Behav Assess 384: 528–537

7. Grills AE, Ollendick TH (2003) Multiple informant agreement and the anxiety disorders interview schedule for parents and children. J Am Acad Child Adolesc Psychiatry 42: 30–40 [PubMed: 12500074]

8. Dirks MA, Weersing VR, Warnick E, Gonzalez A, Alton M, Dauser C et al. (2014) Parent and youth report of youth anxiety: evidence for measurement invariance. J Child Psychol Psychiatry 55: 284–291 [PubMed: 24552483]

9. Wren FJ, Bridge JA, Birmaher B (2004) Screening for childhood anxiety symptoms in primary care: integrating child and parent reports. J Am Acad Child Adolesc Psychiatry 43: 1364–1371 [PubMed: 15502595]

10. Scaini S, Ogliari A, De Carolis L, Bellodi L, Di Serio C, Brombin C (2017) Evaluation of mother-child agreement and factorial structures of the SCARED questionnaire in an Italian clinical sample. Front Psycol 8: 242

11. Klein RG (1991) Parent-child agreement in clinical assessment of anxiety and other psychopathology: A review. J Anxiety Disord 5: 187–198

12. Cosi S, Canals J, Hernández-Martinez C, Vigil-Colet A (2010) Parent–child agreement in SCARED and its relationship to anxiety symptoms. J Anxiety Disord 24: 129–133 [PubMed: 19864109]

13. Ivarsson T, Skarphedinsson G, Andersson M, Jarbin H (2017) The validity of the Screen for Child Anxiety Related Emotional Disorders Revised (SCARED-R) scale and subscales in Swedish youth. Child Psychiatry Hum Dev 49: 1–10

14. Jensen PS, Xenakis SN, Davis H, Degroot J (1988) Child psychopathology rating scales and interrater agreement: Child and family characteristics. J Am Acad Child Adolesc Psychiatry 27: 451–461 [PubMed: 3182601]

15. Dimitrov DM (2010) Testing for factorial invariance in the context of construct validation. Meas Eval Couns Dev 43: 121–149

16. van de Schoot R, Lugtig P, Hox J (2012) A checklist for testing measurement invariance. Eur J Dev Psychol 9: 486–492

17. Boyd RC, Ginsburg GS, Lambert SF, Cooley MR, Campbell KD (2003). Screen for Child Anxiety Related Emotional Disorders (SCARED): psychometric properties in an African-American parochial high school sample. J Am Acad Child Adolesc Psychiatry 42: 1188–1196 [PubMed: 14560168]

18. Arab A, El Keshky M, Hadwin JA (2016) Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED) in a non-clinical sample of children and adolescents in Saudi Arabia. Child Psychiatry Hum Dev 47: 554–562 [PubMed: 26424720]

19. Su L, Wang K, Fan F, Su Y, Gao X (2008) Reliability and validity of the screen for child anxiety related emotional disorders (SCARED) in Chinese children. J Anxiety Disord 22: 612–621 [PubMed: 17628391]

20. Isolan L, Salum GA, Osowski AT, Amaro E, Manfro GG (2011) Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED) in Brazilian children and adolescents. J Anxiety Disord 25: 741–748 [PubMed: 21514788]

21. Kaufman J, Birmaher B, Brent D, Rao UMA, Flynn C, Moreci P et al. (1997) Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. J Am Acad Child Adolesc Psychiatry 36: 980–988 [PubMed: 9204677]

22. Hale WW, Raaijmakers Q, Muris P, Meeus WIM (2005) Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED) in the general adolescent population. J Am Acad Child Adolesc Psychiatry 44: 283–290 [PubMed: 15725973]

23. Research Units on Pediatric Psychopharmacology (RUPP) Anxiety Study Group (2002) The pediatric anxiety rating scale (PARS): development and psychometric properties. J Am Acad Child Adolesc Psychiatry 41: 1061–1069 [PubMed: 12218427]

24. Birmaher B, Axelson D, Monk K, Kalas C, Clark D, Ehmann M et al. (2003) Fluoxetine for the treatment of anxiety disorders. J Am Acad Child Adolesc Psychiatry 42: 415–423 [PubMed: 12649628]

25. Walkup JT, Albano AM, Piacentini J, Birmaher B, Compton SN, Sherrill JT et al. (2008) Cognitive behavioral therapy, sertraline, or a combination on childhood anxiety. N Engl J Med 359: 2753–66 [PubMed: 18974308]

26. Silk J, Tan P, Ladouceur C, Meller S, Siegle G, McMakin D et al. (2016) A randomized clinical trial comparing individual cognitive behavioral therapy and child-centered therapy for child anxiety disorders. J Clin Child Adolesc Psychol, 1–13. Retrieved from 10.1080/15374416.2016.11384

27. Wechsler D (1999) Wechsler abbreviated scale of intelligence. Pearson

28. Dwyer SB, Nicholson JM, Battistutta D (2003) Population level assessment of the family risk factors related to the onset or persistence of children's mental health problems. J Child Psychol Psychiatry 44: 699–711 [PubMed: 12831114]

29. Muthén LK, & Muthén BO (2010). Mplus: Statistical analysis with latent variables: User's guide (pp. 1998–2007). Los Angeles: Muthén & Muthén.

30. Gomez R, Vance A, Gomez RM (2014) The factor structure of anxiety and depressive disorders in a sample of clinic-referred adolescents. J Abnorm Child Psychol 42: 321–332 [PubMed: 23942827]

31. Muthén B, & Asparouhov T (2002). Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus. Mplus web notes, 4(5), 1–22.

32. Cheung GW & Rensvold RB (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Modeling 9: 233–255

33. Hu LT, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Eq Modeling 6: 1–55

34. Schermelleh-Engel K, Moosbrugger H, Muller H (2003) Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. Psychol Res 8: 23–74

35. Revelle W (2011). Psych: Procedures for personality and psychological research. Evanston: Northwestern University

36. Team RC (2013). R: A language and environment for statistical computing

37. Bates D, Maechler M, Bolker B, Walker S(2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67, 1–48

38. Hariz N, Bawab S, Atwi M, Tavitian L, Zeinoun P, Khani M, et al. (2013) Reliability and validity of the Arabic Screen for Child Anxiety Related Emotional Disorders (SCARED) in a clinical sample. Psychiatry Res 209: 222–228 [PubMed: 23312477]

39. Bauer DJ (2017) A more general model for testing measurement invariance and differential item functioning. Psychol Methods 22: 507 [PubMed: 27266798]

40. Cicchetti DV (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 6: 284

41. Nunnally JC (1978): Psychometric Theory. McGraw-Hill, New York

42. Jansen M, Bodden DH, Muris P, van Doorn M, Granic I (2017) Measuring anxiety in children: the importance of separate mother and father reports. Child Youth Care Forum 46: 643–659 [PubMed: 28989266]

43. Gonzalez A, Weersing VR, Warnick E, Scahill L, Woolsten J (2012) Cross-ethnic measurement equivalence of the SCARED in an outpatient sample of African American and non-Hispanic white youths and parents. J Am Acad Child Adolesc Psychiatry 41: 361–369

44. Wren FJ, Berg EA, Heiden LA, Kinnamon CJ, Ohlson LA, Bridge JA et al. (2007) Childhood anxiety in a diverse primary care population: parent child reports, ethnicity and SCARED factor structure. J Am Acad Child Adolesc Psychiatry 46: 332–340 [PubMed: 17314719]

Table 1:

Sample Characteristics

	Total	Healthy	Anxious
N (% of total)	1,092	635 (58.15%)	457 (41.85%)
Age***	12.52 (2.60)	12.88 (2.47)	12.02 (2.70)
Sex (n / % female)	553 (50.60%)	326 (51.30%)	227 (49.70%)
Ethnicity (n / % white)	706 (64.70%)	409 (64.40%)	297 (65.00%)
Highest parental education[a] *	1.74 (.91)	1.68 (.80)	1.80 (0.99)
Household income[a]	7.41 (1.63)	7.49 (1.50)	7.33 (1.75)
IQ[b]***	111.84 (13.18)	110.73 (11.90)	113.22 (14.52)
FRFC[c] ***	5.07 (3.50)	3.51 (2.91)	6.45 (3.40)
SCARED-C***	17.33 (12.56)	12.65 (9.37)	23.84 (13.51)
SCARED-P***	14.95 (13.82)	12.09 (12.36)	18.91 (14.74)
SCARED-P/C mean***	16.14 (10.16)	12.37 (7.94)	21.37 (7.94)
RDS***	−2.39 (16.88)	−0.55 (15.12)	−4.94 (18.78)
absRDS***	13.01 (11.00)	11.23 (10.13)	15.50 (11.67)
SMDS***	0 (1.28)	0.17 (1.13)	−0.23 (1.42)
absSMDS***	0.96 (0.84)	0.84 (0.78)	1.14 (0.88)

Demographic and clinical characteristics are displayed here for the full sample and for the healthy and anxious subsamples. Higher raw discrepancy scores (RDS) and standardized mean difference scores (SMDS) indicate higher scores for parents compared to their child. Higher absolute RDS (absRDS) and SMDS (absSMDS) indicate greater discrepancy regardless of the direction of discrepancy. Significant group differences between the healthy and anxious subsamples are noted

*
$p < .05$,

**
$p < .01$,

***
$p < .001$.

[a]
Information on highest parental education and household income was collected from $n = 798$ participants. These variables were assessed on ordinal scales: Highest parental education (1 = graduate professional degree, 2 = standard college graduation, 3 = partial college, 4 = high school graduation, 5 = partial high school, 6 = junior high school, 7 = less than 7 years of school); Household income (1 = Under $5,000, 2 = $5,000 – $9,999, 3 = $10,000 – $14,999, 4 = $15,000 – $24,999, 5 = $25,000 – $39,999, 6 = $40,000 – $59,999, 7 = $60,000 – $89,999, 8 = $90,000 – $179,999, 9 = Over $180,000).

[b]
IQ scores were collected from $n = 974$ participants.

[c]
FRFC scores were collected from $n = 358$ participants.

Table 2:

Parent-Child Agreement

	Total	Healthy	Anxious
SCARED- Total	.17***	.05	.09*
SCARED- GAD subscale	.16***	.06	.11*
SCARED- Social phobia subscale	.19***	.13***	.18***
SCARED- Panic subscale	.14***	.02	.08
SCARED- School anxiety subscale	.15***	.05	.14**
SCARED- Separation subscale	.17***	.04	.16***

Intra-class correlation values are presented for parent-child scores.

*
$p < .05$,

**
$p < .01$,

***
$p < .001$.

Table 3:

Correlates of the SCARED

Correlates	RDS	absRDS	SMDS	absSMDS
Age	$r = 0.01$	$r = -0.06$	$r = 0.02$	$r = -0.06$ [*]
Sex	$t = 1.94$	$t = 0.37$	$t = 2.03$ [*]	$t = 0.92$
Ethnicity	$t = -1.04$	$t = 1.88$	$t = -1.09$	$t = 1.77$
Highest parental education[a]	$r = -0.15$ [***]	$r = 0.06$	$r = -0.16$ [***]	$r = 0.03$
Income[a]	$r = 0.09$ [*]	$r = -0.02$	$r = 0.09$ [*]	$r = -0.02$
IQ[b]	$r = 0.07$ [*]	$r = 0.01$	$r = 0.07$ [*]	$r = 0.02$
FRFC[c]	$r = -0.06$	$r = 0.07$	$r = -0.09$	$r = 0.08$
Diagnosis	$t = -4.12$ [***]	$t = 6.30$ [***]	$t = -5.00$ [***]	$t = 5.78$ [***]
SCARED-P/C mean	$r = 0.07$ [*]	$r = 0.46$ [**]	$r = 0.02$	$r = 0.46$ [**]

Correlations between SCARED discrepancy scores are presented (r = Spearman's rho; t = independent-samples t-test)

[*]
$p < .05$,

[**]
$p < .01$,

[***]
$p < .001$. Higher raw discrepancy scores (RDS) and standardized mean difference scores (SMDS) indicate higher scores for parents compared to their child. Higher absolute RDS (absRDS) and SMDS (absSMDS) indicate greater discrepancy regardless of the direction of discrepancy. The direction of t-test results were: sex (male > female), ethnicity (non-white > white), and diagnosis (anxious > healthy).

[a]
Information on highest parental education and household income was collected from $n = 803$ participants. Note that a value of 1 indicates the highest and 7 indicates the lowest parental educational attainment.

[b]
IQ scores were collected from $n = 974$ participants.

[c]
FRFC scores were collected from $n = 358$ participants.

Table 4:

Regression Analyses

Variable	RDS		absRDS		SMDS		absSMDS	
	B	*t*	*B*	*t*	*B*	*t*	*B*	*t*
Age	0.01	0.03	−0.11	−0.73	0	0.03	−0.01	−0.80
Sex	2.18	1.72	0.46	0.60	0.17	1.72	0.07	1.21
Ethnicity	−0.04	−0.03	0.29	0.33	0	−0.03	0.03	0.50
Highest parental education	−2.83	−3.35 ***	0.75	1.45	−0.22	−3.35 ***	0.02	0.64
Income	0.06	−0.20	0.10	0.38	−0.01	−0.20	0	0.04
IQ	−0.09	−0.90	0	0	0	−0.43	0	0.22
SCARED- P/C mean	0.37	5.42 ***	0.45	10.93 ***	0.02	4.01 ***	0.04	11.87 ***
Diagnosis	−8.00	−5.51 ***	−0.51	−0.58	−0.61	−5.51 ***	−0.09	−1.32
Model R²		0.08		0.18		0.07		0.19
Model F		8.12		19.31		7.22		21.5

Higher raw discrepancy scores (RDS) and standardized mean difference scores (SMDS) indicate higher scores for parents compared to their child. Higher absolute RDS (absRDS) and SMDS (absSMDS) indicate greater discrepancy regardless of the direction of discrepancy. Binary variables were coded as sex (male > female), ethnicity (non-white > white), and diagnosis (anxious > healthy). Information on age, sex, ethnicity, highest parental education, income, child IQ, SCARED-P/C mean score, and diagnosis was available in an overlapping sample of $n = 730$ dyads.

[*]
$p < .05$,

[**]
$p < .01$,

[***]
$p < .001$.

Table 5:

Within-Dyad Parent-Child Invariance Test

Invariance Level	χ^2	df	χ^2	CFI	RMSEA	Lower CI	Upper CI	Change in CFI	Change in RMSEA
Configural	5442.399	3194	--	**0.966**	**0.025**	0.024	0.027	--	--
Metric	5609.782	3230	170.44 ***	**0.964**	**0.026**	0.025	0.027	**−0.002**	**−0.001**
Threshold	5976.285	3307	1028.49 ***	**0.960**	**0.027**	0.026	0.028	**−0.004**	**0.001**
Residual	5752.553	3266	241.29 ***	**0.963**	**0.026**	0.025	0.028	**0.003**	**−0.001**

Within-dyad invariance tests using a mean- and variance-adjusted weight least squares (WLSMV) estimator are presented here. Good model fit (noted in bold) was established by a comparative fit index (CFI) > .95 and root mean square error of approximation (RMSEA) < .06. Measurement invariance at each level (noted in bold) was established by a decrement in CFI < .01 and an increase in RMSEA < .015.

***
$p < .001$.

Table 6:

Invariance Tests by Age Group

Child Report

Invariance Level	χ^2	df	χ^2	CFI	RMSEA	Lower CI	Upper CI	Change in CFI	Change in RMSEA
Configural	3286.088	1538	--	0.967	0.046	0.044	0.048	--	--
Metric	2962.415	1579	54.17	0.974	0.040	0.038	0.042	0.007	−0.006
Threshold	3073.168	1661	159.77[***]	0.974	0.040	0.037	0.042	0	0
Residual	3272.629	1610	97.50[***]	0.969	0.044	0.041	0.046	−0.005	0.004

Parent Report

Invariance Level	χ^2	df	χ^2	CFI	RMSEA	Lower CI	Upper CI	Change in CFI	Change in RMSEA
Configural	2780.89	1538	--	0.954	0.039	0.036	0.041	--	--
Metric	2728.21	1579	103.82[***]	0.958	0.037	0.034	0.039	0.004	−0.002
Threshold	3159.85	1661	168.31[***]	0.945	0.041	0.039	0.043	−0.013	0.004
Residual	3159.85	1660	91.42[***]	0.946	0.041	0.039	0.043	.001	0

Invariance tests using a mean- and variance-adjusted weight least squares (WLSMV) estimator comparing older and younger children on child-report (top panel) and parent-report (bottom panel) are presented here. Good model fit (noted in bold) was established by a comparative fit index (CFI) > .95 and root mean square error of approximation (RMSEA) < .06. Measurement invariance at each level (noted in bold) was established by a decrement in CFI < .01 and an increase in RMSEA < .015.

[***] $p < .001$.

Table 7:

Test-etest Reliability of SCARED Scores

	Parent-Report	Child-Report	Discrepancy (SMDS)
Total	.86	.62	.66
Social subscale	.85	.60	.59
GAD subscale	.85	.62	.62
SAD subscale	.85	.59	.59
Panic subscale	.74	.61	.64
School subscale	.79	.60	.56

Test-retest reliability values are presented here, indicated by intra-class correlation values from linear mixed effects models.