HHS Public Access

Author manuscript

Behav Ther. Author manuscript; available in PMC 2021 March 01.

Published in final edited form as:

Behav Ther. 2020 March ; 51(2): 283–293. doi:10.1016/j.beth.2019.10.005.

The Clinician Affective Reactivity Index: Validity and Reliability of a Clinician-Rated Assessment of Irritability

Simone P. Haller, **Katharina Kircanski**, **Argyris Stringaris**, **Michal Clayton**, **Hong Bui**, **Courtney Agorsor**, **Sofia I. Cardenas**, **Kenneth E. Towbin**, **Daniel S. Pine**, **Ellen Leibenluft**, **Melissa A. Brotman**

Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health

Abstract

Irritability is impairing in youth and is the core feature of disruptive mood dysregulation disorder (DMDD). Currently, there are no established clinician-rated instruments to assess irritability in pediatric research and clinical settings. Clinician-rated measures ensure consistency of assessment across patients and are important specifically for treatment research. Here, we present data on the psychometric properties of the Clinician Affective Reactivity Index (CL-ARI), the first semistructured interview focused on pediatric irritability. The CL-ARI was administered to a transdiagnostic sample of 98 youth (*M* age = 12.66, *SD* = 2.47; 41% female). With respect to convergent validity, CL-ARI scores were (a) significantly higher for youth with DMDD than for any other diagnostic group, and (b) showed uniquely strong associations with other clinician-, parent-, and youth-report measures of irritability compared to measures of related constructs, such as anxiety. The three subscales of the CL-ARI (temper outbursts, irritable mood, impairment) showed excellent internal consistency. Test-retest reliability of the CL-ARI was adequate. These data support that irritability can be feasibly, validly, and reliably assessed by clinicians using the CL-ARI. A validated, gold-standard assessment of pediatric irritability is critical in advancing research and treatment efforts.

Keywords

irritability; measurement; clinician; validity; reliability

PREVALENCE RATES OF CHRONIC, severe irritability range from 0.1%–5.3% (Brotman et al., 2006; Copeland, Angold, Costello, & Egger, 2013). Chronic, severe irritability is impairing in youth, associated with significant family burden, and linked to socioeconomic underachievement and trajectories of affective psychopathology in adulthood (Brotman,

Address correspondence to Simone P. Haller, Ph.D., Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD, 20892; simone.haller@nih.gov.

Conflict of Interest Statement

The authors declare that there are no conflicts of interest.

Kircanski, Stringaris, Pine, & Leibenluft, 2017; Leibenluft, 2017; Stringaris, Cohen, Pine, & Leibenluft, 2009; Vidal-Ribas, Brotman, Valdivieso, Leibenluft, & Stringaris, 2016). Disruptive mood dysregulation disorder (DMDD) was added to the DSM-5 in recognition of the clinical significance of chronic, severe irritability (American Psychiatric Association, 2013). Currently, however, there are no clinician-rated instruments to assess irritability in pediatric research and clinical settings (Brotman, Kircanski, & Leibenluft, 2017). Such measures are essential for evaluating treatment efficacy. Here, we describe the Clinician Affective Reactivity Index (CL-ARI), a clinician-rated index of irritability. We present data on its psychometric properties, including criterion validity, internal consistency, and test-retest reliability, to support the use of this new instrument.

The CL-ARI is a semistructured, parent-child interview assessing irritability over the past week. Irritability in the CL-ARI is conceptualized around two clinically important components: temper outbursts and prolonged irritable mood. Additionally, the CL-ARI assesses irritability-related impairment. The three components are rated on separate subscales. The first subscale assesses *temper outbursts*, developmentally inappropriate behavioral manifestations of anger that are disproportionate in intensity and duration to the situation. Specifically, the frequency and duration of temper outbursts are quantified across three different levels of severity: mild (e.g., snapping or mild arguing), moderate (e.g., yelling, verbally threatening someone or misusing property), and severe (e.g., shoving or kicking another person, or repeatedly engaging in behaviors that destroy property). Thus, even mild outbursts are reflected in the scoring to capture the full spectrum of irritability. *Irritable mood*, assessed in the second subscale, describes the presence of persistent negative affect (e.g., general level of grumpiness, crankiness). Items in this irritable mood subscale capture the frequency, duration, and severity of irritable mood. The third subscale covers *impairment* (i.e., the extent to which the child's irritability affects functioning). Items require the parent and child to describe impairment in family, peer, and school settings.

Previous instruments have been developed to assess DMDD symptoms, but current clinician-rated instruments are confined to the diagnostic criteria. The DMDD module developed by Leibenluft and colleagues (Wiggins et al., 2016) for use with the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime Version (K-SADS-PL; Kaufman et al., 1997) has been used to establish the diagnosis of DMDD. This includes differentiating DMDD from narrow phenotype bipolar disorder, which can present with circumscribed episodic irritability (Brotman et al., 2007). As irritability presents in other psychopathology, such as oppositional-defiant disorder (ODD), attention-deficit/hyperactivity disorder (ADHD), anxiety disorders, major depressive disorder (MDD), and autism spectrum disorders, other K-SADS modules also contain individual items querying irritability. However, no K-SADS module surveys irritability as a cross-disorder dimension. Indeed, the clinical significance of irritability in many diagnoses indicates the need for such a cross-disorder measure (Cornacchio, Crum, Coxe, Pincus, & Comer, 2016; Dickstein et al., 2009; Eyre et al., 2017; Haller et al., 2018; Karalunas, Gustafsson, Fair, Musser, & Nigg, 2019; Stoddard et al., 2016; Stoddard et al., 2014; Stringaris, Maughan, Copeland, Costello, & Angold, 2013; Waxmonsky et al., 2013).

Previous research on irritability indices has relied primarily on parent- and youth-report rating scales. Despite their utility, such measures have limitations (Anastasi & Urbina, 1997). In particular, in many psychosocial treatment studies, parents and children are not blind to the treatment they are receiving, creating a need for clinician-rated measures. The youth- and parent-reported Affective Reactivity Index (ARI; Stringaris et al., 2012) provides a guide for generating such a measure. This screening tool consists of six symptom items (temper outbursts and irritable mood) and one impairment item. However, by design, the youth- and parent-report ARI does not query details of temper outbursts and irritable mood. Building on this, the CL-ARI requires parents and children to describe the frequency, duration, and severity of individual temper outbursts and irritable mood.

We examined the psychometric properties of the CL-ARI. The CL-ARI is designed to cover these different components of irritability in detail and as a flexible tool capable of assessing irritability across diagnoses. As a transdiagnostic tool, the CL-ARI should be able to differentiate irritability across distinct clinical groups. Clinical groups in this study included youth with primary DMDD, "subthreshold DMDD," ADHD, and anxiety disorders, and youth with no diagnosis. There has been controversy surrounding the exact threshold criteria for DMDD, with current criteria possibly too high and failing to capture many children impaired by irritability (Althoff et al., 2016; Axelson et al., 2012; Freeman, Youngstrom, Youngstrom, & Findling, 2016; Fristad et al., 2016; Margulies, Weintraub, Basile, Grover, & Carlson, 2012; Mayes et al., 2015; Mayes, Waxmonsky, Calhoun, & Bixler, 2016; Mitchell et al., 2016; Roy, Lopes, & Klein, 2014; Stringaris, 2011, 2013; Tufan et al., 2016). In the current study, we included youth with "subthreshold DMDD," which represents a research designation designed to capture youth with significant irritability that fell just short of diagnostic criteria (e.g., exhibited either significant temper outbursts or irritable mood, as well as impairment in at least one setting). First, we aimed to evaluate criterion validity of the CL-ARI, using as relevant metrics (a) youth's primary diagnoses, (b) scores on another, more limited clinician-rated assessment of illness severity (CGI-S for DMDD, Clinical Global Impressions of Scale-Severity), and (c) scores on parent- and youth-report measures of irritability (i.e., ARI-P and ARI-C). We expected that scores on the CL-ARI would be significantly higher for youth with DMDD than for other diagnostic groups, and that CL-ARI scores would show uniquely strong associations with other clinician-, parent-, and youth-report measures of irritability compared to measures of related constructs, such as anxiety or depression. Second, we aimed to assess two forms of reliability, internal consistency, and test-retest reliability, of the CL-ARI. We expected the CL-ARI to show strong psychometric properties including good criterion validity, internal consistency, and test-retest reliability.

Method

PARTICIPANTS

Participants were 98 youths ages 7 to 17 years (M = 12.66 years, SD = 2.47; 41% female), with primary diagnoses of DMDD, subthreshold DMDD (i.e., child met criteria for temper outbursts or irritable mood, as well as impairment in at least one setting), ADHD, or anxiety disorders (generalized, social, and/or separation anxiety disorder), or no diagnosis.

Diagnoses were established using the K-SADS-PL (Kaufman et al., 1997), including an additional extended DMDD module (Wiggins et al., 2016). Table 1 presents demographic and clinical characteristics of the sample. For details on medications, see Table S1 in the supplementary material.

Participants were recruited through the NIMH Emotion and Development Branch for characterization and treatment research. An initial evaluation visit included a diagnostic interview to determine eligibility for different research protocols recruiting various clinical groups based on a primary symptomatic target. Specifically, the clinical groups included ADHD, DMDD, subthreshold DMDD, a primary anxiety disorder, and healthy comparison participants with no current or lifetime diagnosis. Exclusionary criteria included the following: cardinal bipolar symptoms, met criteria for schizophrenia, schizophreniform disorder, schizoaffective illness, pervasive developmental disorder, or posttraumatic stress disorder, alcohol or substance abuse within the last 3 months, and an IQ below 70 as assessed by the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). Parents provided informed consent and youth provided assent in accordance with NIMH Institutional Review Board approval.

MEASURES

CL-ARI

Administration.: The CL-ARI assesses irritability over a 1-week period, although a version covering a shorter time frame (3 days) is available in identical format. The CL-ARI is available upon request from the senior author.

The CL-ARI total score is the weighted sum (see Scoring below) of three subscale scores: temper outbursts, irritable mood between outbursts, and impairment. The CL-ARI is designed to be administered by a trained clinician (master's level or above) to parents and children. As a semistructured interview, the CL-ARI contains verbal prompts and examples to lead the administrator through the items. A total of 12 items query the *frequency, duration,* and *severity* of mild, moderate, and severe temper outbursts and irritable mood between outbursts, as well as functional impairment in home, school, and peer settings. Parent and children are interviewed separately; the clinician synthesizes information into a consensus rating. Items are scored on Likert scales. Temper outburst frequency is scored on a 5-point scale, separately for mild, moderate, and severe outbursts (0 = *none*, 4 = *more than one outburst every day*). Temper outburst duration is scored on a 6-point scale, again separately for mild, moderate, and severe outbursts (0 = *none*, 5 = *>60 min*). The irritable mood frequency item has a 4-point scale (0 = *none*, 3 = *four or more days*). The irritable mood duration item is only completed if the irritable mood frequency rating is scored at the maximum score (3 = *four or more days*). The mood duration item queries whether irritable mood was present for most of the day (i.e., at least half the day with a *yes/no* response option). The irritable mood severity item has a 6-point scale (0 = *not present*, 5 = *severe*). Last, impairment is rated separately in three different settings (family, school, peers) on 6-point scales (0 = *none*, 5 = *severe*).

Scoring.: The range of possible subscale scores is as follows: temper outbursts (0–27); mood (0–8); impairment (0–15). Given the nascent status of irritability work, total scores weight the three subscale scores equally. That is, we did not make *a priori* assumptions that one component is more important than the others in the overall phenotype. To aid interpretation of total scores, each subscale score is divided by the total possible score of the subscale; these proportional subscale scores are then summed, divided by the number of subscales (3), and multiplied by 100.

Total score = [(temper observed score/total possible subscale score [27])+
(mood observed score/total possible subscale score [8])+(impairment observed score/total possible subscale score [15])/3] * 100

Thus, subscales are weighted equally in the total score, which ranges from 0–100.

Other Clinician-Rated Measures: The following measures, widely used in clinical research, were administered to participants:

Clinical Global Impressions Scale - Severity of DMDD (CGI-S DMDD; Dickstein et al., 2009; Stoddard et al., 2016; Towbin et al., 2019; Walkup et al., 2001) is a single-item, 7-point measure of overall DMDD severity (1 = *normal*, 7 = *most extreme*) over the past month.

Children's Global Assessment Scale (CGAS; Bird, Canino, Rubio-Stipec, & Ribera, 1987; Shaffer et al., 1983) measures global impairment in functioning. The measure assesses function relative to peers over the past month (1 = lowest, 100 = highest) and has strong psychometric properties (Green, Shirk, Hanze, & Wanstrath, 1994).

Pediatric Anxiety Rating Scale (PARS, 2002) assesses the frequency and severity of symptoms common to pediatric anxiety disorders, including separation, social, and generalized anxiety symptoms over the past week. The PARS is widely used in clinical anxiety research with established psychometric properties (PARS, 2002).

Children's Depression Rating Scale (CDRS; Poznanski et al., 1984) is a clinician-rated depression screening tool that assesses 17 symptom areas of depression over the past week. It has been used widely with excellent psychometrics (Poznanski et al., 1984).

Parent- and Youth-Report Measures: *Affective Reactivity Index* (ARI; Stringaris et al., 2012) is a concise, 7-item parallel parent-report (ARI-P) and child-report (ARI-C) questionnaire, which surveys irritable mood and outbursts over the past 6 months. It has been validated in multiple clinical samples with good psychometric properties (Stringaris et al., 2012).

Screen for Child Anxiety Related Emotional Disorders (SCARED; Birmaher et al., 1997) is a 38-item questionnaire with both parent-report (SCARED-P) and child-report (SCARED-C) forms. It assesses symptoms associated with each of the major child anxiety disorder diagnoses over the past 3 months. The SCARED parent- and child-report questionnaire is

widely used in clinical and research settings with good psychometric properties (Behrens, Swetlitz, Pine, & Pagliaccio, 2019).

Children's Depression Inventory (CDI; Kovacs, 1981) is a 27-item youth-report rating scale of depressive symptoms in children over the past two weeks. Widely used in clinical research, it has well-established validity with sound psychometric properties (Saylor, Finch, Spirito, & Bennett, 1984).

PROCEDURES

A comprehensive pre-study evaluation was conducted, which included the administration of the K-SADS-PL. Clinician-rated measures, including the CL-ARI, were completed on a separate visit or over the phone. The same clinician completed ratings with both parent and child informants. Four master's or doctoral-level clinicians were trained to administer the CL-ARI. To use the instrument, clinicians had to achieve an interrater reliability criterion of ICC = .90 across $N = 10$ cases. Given the semistructured nature of the measure, administration time varied depending on the degree of irritability symptoms reported, as well as the reporting styles of child and parent. The total administration time (i.e., both child and parent interviews) for participants diagnosed with DMDD was approximately 20–30 minutes and for healthy volunteers was approximately 5–10 minutes. All clinician- and youth-report measures were completed within three months of the CL-ARI.

A subset of 26 youth ages 9 to 16 (M age = 12.64, $SD = 2.11$) with primary DMDD ($n = 24$) or subthreshold DMDD ($n = 2$) completed the CL-ARI twice. Re-test interviews were completed one week after the initial interview.

STATISTICAL ANALYSES

Complete CL-ARI data were available for all participants. Ninety participants (91.8%) completed all clinician ratings (CGI-S, CGAS, PARS, CDRS) within the 3-month time span. Seventy-one participants (72.5%) completed all youth-/parent-report questionnaires within the 3-month time span. The CDI was completed by a smaller subsample ($n = 45$, 45.9%). Missing single item scores on questionnaires (<1% missing) were imputed using the participant's mean item score on that questionnaire. Participants with missing data were excluded listwise from analyses examining clinician-, youth-, or parent-report measures.

Validity

Repeated-measures analyses of variance (ANOVAs) were used to examine average frequency and duration of outbursts, mood severity and frequency, and impairment across the full sample. One-way ANOVAs were used to examine CL-ARI scores across diagnostic groups. Youth with primary anxiety disorder were excluded from this statistical comparison due to low power, with $n = 7$ in this group. When examining other properties of the CL-ARI, the sample was collapsed across diagnostic categories. Pearson correlation coefficients were used to examine the relations between CL-ARI scores and the other clinical measures. Fisher's r-to-z transformations (Fisher, 1921; Steiger, 1980) were used to test for differences between these correlation coefficients.

Reliability

Cronbach's α was used to measure internal consistency. Test-retest reliability data were analyzed using intraclass correlation coefficients (ICC[2,1]); McGraw & Wong, 1996; Shrout & Fleiss, 1979).

Results

DESCRIPTIVE STATISTICS

Table 2 presents means and standard deviations of CL-ARI item scores, both for youth in the DMDD group specifically and for the full sample. As shown by the item scores in the full sample, frequency of outbursts fell on a spectrum in which mild outbursts (e.g., snapping, mild arguing, name calling or huffing and puffing) were most frequent. Mild outbursts were reported as occurring every day of the week [$F(1.74, 168.55) = 154.59$, $p<.001$, $\eta^2 = .61$, mild-moderate: $t(97) = 9.72$, $p<.001$, mild-severe: $t(97) = 15.78$, $p<.001$]. Moderate outbursts (e.g., verbal threats, physical displays of anger such as angrily misusing property) were reported significantly less frequently than mild outbursts, on average, 1-2/3 or more times a week (moderate-severe: $t[97] = 8.92$, $p<.001$). Severe outbursts (e.g., physically pushing, shoving, kicking another person) were relatively rare in our sample. Either none were reported, or they were reported at a frequency of 1–2 times a week. Mild and moderate outbursts both lasted on average a few minutes [$F(2, 194) = 52.12$, $p<.001$ $\eta^2 = .35$, mild-moderate: $t(97) = -1.3$, $p = .19$], whereas severe outbursts were significantly shorter in duration [mild-severe: $t(97) = 5.52$, $p<.001$, moderate-severe: $t(91) = 4.56$, $p<.001$]. Across the full sample, irritable mood was present on average 1–3 days a week, at mild to moderate severity. Irritability-related impairment was scored greatest in the family setting [$F(2, 194) = 19.04$, $p<.001$ $\eta^2 = .16$, family-peers: $t(97) = 8.59$, $p<.001$, family-school: $t(97) = 8.22$, $p<.001$, school-peers: $t(97) = -1.20$, $p = .23$]. For further analyses examining relations among items and subscales, see supplementary material.

With respect to demographic variables, age was negatively correlated with CL-ARI total scores [$r(96) = -.42$, $p<.001$] and subscale scores of temper outbursts [$r(96) = -.46$, $p<.001$], irritable mood [$r(96) = -.39$, $p<.001$] and impairment [$r(96) = -.29$, $p<.001$]. Hence, younger participants had significantly higher scores on CL-ARI items than older participants. However, CL-ARI total and subscale scores did not differ significantly by sex (all ts<1.4, ps>.17).

Validity

Primary diagnosis.—Figure 1 illustrates the distribution of CL-ARI total scores by primary diagnosis in the full sample. A one-way ANOVA indicated that CL-ARI total scores differed significantly by primary diagnosis, $F(3,87) = 47.80$, $p<.001$. Specifically, Scheffe post-hoc tests showed that youth in the DMDD ($M = 48.98$, $SD = 14.32$) and subthreshold DMDD ($M = 39.43$, $SD = 11.48$) groups scored significantly higher on the CL-ARI than youth in the ADHD ($M = 18.24$, $SD = 14.35$; both ps<.001) and healthy groups ($M = 0.69$, $SD = 2.06$; both ps<.001). In addition, the ADHD group had significantly higher scores than the healthy group ($p = .019$). For subscale scores by primary diagnosis and associated statistics, see Figure S2 in the supplementary material.

Clinician-rated measures.—Table 3 summarizes correlations between CL-ARI total scores and scores on other clinician-, parent-, and youth-report measures of irritability in the full sample. As shown, the correlation between CL-ARI total scores and clinician-rated DMDD severity (CGI-S DMDD) was high, $r(88) = .89$, $p<.001$. The association between CL-ARI total scores and clinician-rated global impairment (CGAS) was also high, with increasing CL-ARI total scores associated with decreased functioning ($r(88) = -.75$, $p<.001$). Correlations with clinician-rated measures of other dimensions of psychopathology were moderate [PARS: $r(88) = .51$, $p<.001$; CDRS: $r(88) = .52$, $p<.001$]. As expected, Fisher's r-to-z transformations indicated that the correlation of CL-ARI total scores with CGI-S DMDD ratings was significantly stronger than the correlations of CL-ARI total scores with ratings of anxiety (z-score: 6.55, $p<.001$) and depression (z-score: 6.53, $p<.001$).

Parent- and youth-report measures.—As shown in Table 3, correlations with parent- and youth-report measures of irritability were moderate to high [ARI-P: $r(69) = .68$, $p<.001$; ARI-C: $r(69) = .42$, $p<.001$]. Correlations with CL-ARI total scores were stronger for ARI-P than ARI-C (z-score = 2.63, $p = .004$), suggesting that either greater weight is given to parents' reports or that there is greater consistency in parent-report across measures. Correlations with measures of other dimensions of psychopathology were moderate to low [SCARED-P: $r(69) = .32$, $p<.01$, SCARED-C: $r(69) = .09$, $p>.05$; CDI: $r(43) = .39$, $p<.01$]. As expected, Fisher's r-to-z transformations indicated that CL-ARI total scores were more strongly related to parent-reported irritability than parent-reported anxiety (z-score = 3.12, $p = .001$), and more strongly related to youth-reported irritability than youth-reported anxiety (z-score = 2.56, $p = .005$). For 95% confidence intervals for correlation coefficients between the CL-ARI and clinician- and youth-report measures, please see Table S2 in supplementary material.

Reliability

Internal consistency.—Internal consistency of all items comprising the CL-ARI was good (Cronbach's $\alpha = .89$). Internal consistency for each subscale was equally adequate (temper outbursts: $\alpha = .78$; mood between outbursts: $\alpha = .87$; irritability-related impairment: $\alpha = .75$).

Test-retest reliability.—Test-retest ICCs across a 1-week interval were assessed in the subsample of youth with DMDD or subthreshold DMDD. The ICC for the total score was adequate (ICC = .67, 95% CI [.28,.85]). Also stable across time were subscale scores for temper outbursts (ICC = .61, 95% CI [.14,.83]), mood between outbursts (ICC = .58, 95% CI [.07,.81]), and irritability-related impairment (ICC = .74, 95% CI [.43,.89]).

Discussion

The current study examined the initial psychometric properties of the CL-ARI, the first clinician-rated assessment focused on pediatric irritability. Results indicate that the CL-ARI shows sensitivity to graded levels of irritability associated with different diagnostic categories, good criterion validity, excellent internal consistency, and adequate test-retest

reliability. Thus, the CL-ARI is a promising instrument for use in clinical and research settings.

With respect to demographic characteristics, CL-ARI total and subscale scores decreased with age from 7 to 17. Comparisons between boys and girls revealed no significant differences in total or subscale scores by sex, in line with previous studies in large community samples on irritability prevalence rates (Copeland et al., 2013).

The CL-ARI showed sensitivity to differences in irritability severity across primary diagnoses. As expected, youth with primary DMDD or subthreshold DMDD exhibited the highest CL-ARI total scores. Although youth with DMDD and subthreshold DMDD did not differ in total scores, supplementary analyses indicated that the two groups did differ significantly in subscale scores, specifically on the impairment subscale. CL-ARI scores of youth with ADHD fell in between the DMDD/subthreshold DMDD groups and healthy volunteers. Thus, the CL-ARI was able to capture variations in irritability symptoms that did not reach the diagnostic threshold for DMDD. Youth with DMDD were overrepresented in this sample, with small groups of youths presenting primarily with other diagnoses, specifically anxiety diagnoses. As a characteristic of multiple DSM-5 diagnoses, irritability may increasingly be investigated as a distinct construct across categorical boundaries (Cornacchio et al., 2016; Eyre et al., 2017; Karalunas et al., 2019; Stoddard et al., 2014; Stringaris et al., 2013; Waxmonsky et al., 2013). This initial data suggests that the CL-ARI may be a useful tool for the detailed study of transdiagnostic irritability.

Investigations of criterion validity indicated that the CL-ARI showed uniquely strong associations with other clinician-, parent- and youth-report measures of irritability compared to measures of anxiety and depression. Specifically, the CL-ARI showed the strongest correlation with the measure closest to the construct assessed by the CL-ARI: the CGI-S for DMDD. The CGI-S is a three-item measure designed specifically to measure DMDD severity. Hence, the CGI-S for DMDD is not usually administered to other diagnostic groups and does not provide a detailed assessment of irritability. In contrast to associations with irritability measures, correlations between CL-ARI scores and measures assessing constructs different from (but overlapping with) irritability, such as anxiety and depression, were significantly lower in magnitude. This also attests to the value of measuring irritability as an important, distinct construct in its own right, rather than simply as a severity marker of other diagnoses.

Overall, indices of reliability were adequate. Internal consistency for each subscale was excellent and test re-test reliability was adequate across the subscales and total score. In fact, test-retest coefficients were in a similar range to other clinician-rated measures, such as the PARS (total score ICC = .55; PARS, 2002). Currently, it is unclear to what degree the test-retest coefficients reflect stability of the instrument or natural fluctuations in irritability over the 1-week period. Further evaluations of reliability in future research could combine measures to assess whether changes over time are reported across instruments.

A clinician-rated, detailed assessment tool for irritability is key to advancing the study of this construct. In the current calculations of CL-ARI scores, the three subscales were

weighted equally. However, the relative importance of the temper outburst vs. mood dimension in the overall phenotype remains largely unexplored. Further study of the importance of the different components in the irritability phenotype, and how different components of irritability may differentially characterize youth with different diagnoses (e.g., ADHD, MDD), may lead to a revision of the weighing of subscales.

CLINICAL APPLICATIONS

An advantage of the CL-ARI is that it allows clinicians to assess different components of irritability across diagnostic categories. This is important for treatment, given that treatment may impact temper outburst and mood components differently (Kircanski, Clayton, Leibenluft, & Brotman, 2018). Additionally, a separate assessment of changes in perceived impairment in various settings is possible, which is another important dimension to monitor in treatment trials. Last, the CL-ARI is time-efficient; due to the specific symptom probes and simple scoring, clinicians can be trained to achieve a high criterion inter-rater reliability (ICC = .90).

Importantly, clinicians need to be trained on specific aspects of the instrument. It is critical that clinicians count each reported temper outburst toward only one severity category. For instance, the chronometry of a given temper outburst may be such that it begins as mild (e.g., with mild arguing), but develops into severe behavior (e.g., yelling, slamming doors, and kicking another person). This outburst should only be counted in the severe category, and not also in the mild category.

LIMITATIONS AND FUTURE DIRECTIONS

There are several methodological limitations to the current study. First, clinicians associated with this study were not blind to the child's diagnostic status, nor their own ratings across measures. Thus, correlation coefficients between the CL-ARI and other clinician-rated measures may be inflated, limiting interpretations of associations between scores on different instruments. This methodological limitation specifically impacts the assessment of criterion validity.

Second, the sample of youths presenting primarily with anxiety disorders or ADHD was small. Therefore, analyses examining CL-ARI scores by primary diagnosis need to be interpreted cautiously. Additionally, because of the limited sample size of these groups, examinations of how different aspects of irritability (i.e., mood and temper outburst components) may differentially characterize youths presenting with ADHD or anxiety diagnoses were not possible. Third, as a tool for treatment outcome studies, it would be important to demonstrate that the CL-ARI is sensitive to change.

In the current version of the CL-ARI, parental accommodations to the child's irritability are captured in the impairment subscale. Parental accommodations describe reduced demands placed on the child in order to avoid temper outbursts (e.g., not having to do chores) or in response to a temper outburst (e.g., "giving in" to the temper outburst and letting the child play video games instead of doing chores). This raises the question of how clinicians should rate symptom severity for a child who is not as irritable as s/he otherwise would have been had accommodations not been in place. As accommodations could potentially have large

effects on symptom reports, further study of the role of parental accommodations in attenuating irritability symptoms may lead to the addition of items specifically probing irritability-related accommodations made for the child in different environments.

Another important avenue for future work is to examine the utility of this measure to assess irritability in the context of autism spectrum disorders.

CONCLUSIONS

A clinician-rated, gold-standard measurement tool is critical to refining our understanding of irritability and advancing treatment research. In the current study, we validate the CL-ARI in a research setting. For dissemination efforts, it would be important to gather data on the properties of this tool in diverse clinical settings, in which administrators may not be as highly trained, and patients are not selected based on stringent inclusion/exclusion criteria.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the rating clinicians associated with this project, Mollie Davis, Cheri McNeil, Wanda Wheeler, and Catherine Yokum, for their assistance.

This research was supported by the National Institute of Mental Health Intramural Research Program (NIMH/NIH), ZIAMH002786-13 (Leibenluft), ZIAMH002778-15 (Leibenluft), ZIAMH002781 (Pine), conducted under Clinical Study Protocols 00-M-0198 (ClinicalTrials.gov: NCT00006177), 02-M-0021 (NCT00025935), 01-M-0192 (NCT00018057) and 15-M-0182 (ClinicalTrials.gov: NCT02531893). The funding body was not involved in the study design, in the collection, analysis and interpretation of data, in the writing of the report, or in the decision to submit the article for publication.

References

Althoff RR, Crehan ET, He JP, Burstein M, Hudziak JJ, & Merikangas KR (2016). Disruptive Mood Dysregulation Disorder at Ages 13-18: Results from the National Comorbidity Survey-Adolescent Supplement. Journal of Child and Adolescent Psychopharmacology, 26(2), 107–113. doi:10.1089/cap.2015.0038 [PubMed: 26771536]

American Psychiatric Association. Diagnostic and statistical manual of mental disorders (5th ed). (2013). Washington, DC: Author.

Anastasi A, & Urbina S (1997). Psychological testing (7th ed). Upper Saddle River, NJ: Simon and Schuster.

Axelson D, Findling RL, Fristad MA, Kowatch RA, Youngstrom EA, Horwitz SM, … Birmaher B (2012). Examining the proposed disruptive mood dysregulation disorder diagnosis in children in the Longitudinal Assessment of Manic Symptoms study. Journal of Clinical Psychiatry, 73(10), 1342–1350. doi:10.4088/JCP.12m07674 [PubMed: 23140653]

Behrens B, Swetlitz C, Pine DS, & Pagliaccio D (2019). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Informant Discrepancy, Measurement Invariance, and Test-Retest Reliability. Child Psychiatry and Human Development, 50(3), 473–482. doi:10.1007/s10578-018-0854-0 [PubMed: 30460424]

Bird HR, Canino G, Rubio-Stipec M, & Ribera JC (1987). Further measures of the psychometric properties of the Children's Global Assessment Scale. Archives of General Psychiatry, 44(9), 821–824. doi:10.1001/archpsyc.1987.01800210069011 [PubMed: 3632256]

Birmaher B, Khetarpal S, Brent D, Cully M, Balach L, Kaufman J, & Neer SM (1997). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale construction and psychometric characteristics. Journal of the American Academy of Child & Adolescent Psychiatry, 36(4), 545–553. doi:10.1097/00004583-199704000-00018 [PubMed: 9100430]

Brotman MA, Kassem L, Reising MM, Guyer AE, Dickstein DP, Rich BA, … Leibenluft E (2007). Parental diagnoses in youth with narrow phenotype bipolar disorder or severe mood dysregulation. American Journal of Psychiatry, 164(8), 1238–1241. doi:10.1176/appi.ajp.2007.06101619 [PubMed: 17671287]

Brotman MA, Kircanski K, & Leibenluft E (2017). Irritability in Children and Adolescents. Annual Review of Clinical Psychology, 13, 317–341. doi:10.1146/annurev-clinpsy-032816-044941

Brotman MA, Kircanski K, Stringaris A, Pine DS, & Leibenluft E (2017). Irritability in youths: A translational model. American Journal of Psychiatry, 174(6), 520–532. doi:10.1176/appi.ajp.2016.16070839 [PubMed: 28103715]

Brotman MA, Schmajuk M, Rich BA, Dickstein DP, Guyer AE, Costello EJ, … Leibenluft E (2006). Prevalence, clinical correlates, and longitudinal course of severe mood dysregulation in children. Biological Psychiatry, 60(9), 991–997. doi:10.1016/j.biopsych.2006.08.042 [PubMed: 17056393]

Copeland WE, Angold A, Costello EJ, & Egger H (2013). Prevalence, comorbidity, and correlates of DSM-5 proposed disruptive mood dysregulation disorder. American Journal of Psychiatry, 170(2), 173–179. doi:10.1176/appi.ajp.2012.12010132 [PubMed: 23377638]

Cornacchio D, Crum KI, Coxe S, Pincus DB, & Comer JS (2016). Irritability and Severity of Anxious Symptomatology Among Youth With Anxiety Disorders. Journal of the American Academy of Child & Adolescent Psychiatry, 55(1), 54–61. doi:10.1016/j.jaac.2015.10.007 [PubMed: 26703910]

Dickstein DP, Towbin KE, Van Der Veen JW, Rich BA, Brotman MA, Knopf L, … Leibenluft E (2009). Randomized double-blind placebo-controlled trial of lithium in youths with severe mood dysregulation. Journal of Child and Adolescent Psychopharmacology, 19(1), 61–73. doi:10.1089/cap.2008.044 [PubMed: 19232024]

Eyre O, Langley K, Stringaris A, Leibenluft E, Collishaw S, & Thapar A (2017). Irritability in ADHD: Associations with depression liability. Journal of Affective Disorders, 215, 281–287. doi:10.1016/j.jad.2017.03.050 [PubMed: 28363151]

Fisher RA (1921). On the “Probable Error” of a coefficient of correlation deduced from a small sample. Metron, 1, 3–33.

Freeman AJ, Youngstrom EA, Youngstrom JK, & Findling RL (2016). Disruptive mood dysregulation disorder in a community mental health clinic: Prevalence, comorbidity and correlates. Journal of Child and Adolescent Psychopharmacology, 26(2), 123–130. doi:10.1089/cap.2015.0061 [PubMed: 26745325]

Fristad MA, Wolfson H, Algorta GP, Youngstrom EA, Arnold LE, Birmaher B, … Findling RL (2016). Disruptive mood dysregulation disorder and bipolar disorder not otherwise specified: Fraternal or identical twins? Journal of Child and Adolescent Psychopharmacology, 26(2), 138–146. doi:10.1089/cap.2015.0062 [PubMed: 26859630]

Green B, Shirk S, Hanze D, & Wanstrath J (1994). The Children’s Global Assessment Scale in clinical practice: An empirical evaluation. Journal of the American Academy of Child & Adolescent Psychiatry, 33(8), 1158–1164. doi:10.1097/00004583-199410000-00011 [PubMed: 7982866]

Haller SP, Stoddard J, MacGillivray C, Stiles K, Perhamus G, Penton-Voak IS, … Brotman MA (2018). A double-blind, randomized, placebo-controlled trial of a computer-based Interpretation Bias Training for youth with severe irritability: A study protocol. Trials, 19(1), 626. doi:10.1186/s13063-018-2960-5 [PubMed: 30428909]

Karalunas SL, Gustafsson HC, Fair D, Musser ED, & Nigg JT (2019). Do we need an irritable subtype of ADHD? Replication and extension of a promising temperament profile approach to ADHD subtyping. Psychological Assessment, 31(2), 236–247. doi:10.1037/pas0000664 [PubMed: 30359050]

Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, … Ryan N (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): Initial reliability and validity data. Journal of the American Academy of Child &

Adolescent Psychiatry, 36(7), 980–988. doi:10.1097/00004583-199707000-00021 [PubMed: 9204677]

Kircanski K, Clayton ME, Leibenluft E, & Brotman MA (2018). Psychosocial treatment of irritability in youth. Current Treatment Options in Psychiatry, 5(1), 129–140. doi:10.1007/s40501-018-0141-5 [PubMed: 30319935]

Kovacs M (1981). Rating scales to assess depression in school-aged children. Acta Paedopsychiatrica, 46(5-6), 305–315. [PubMed: 7025571]

Leibenluft E (2017). Pediatric irritability: A systems neuroscience approach. Trends in Cognitive Sciences, 21(4), 277–289. doi:10.1016/j.tics.2017.02.002 [PubMed: 28274677]

Margulies DM, Weintraub S, Basile J, Grover PJ, & Carlson GA (2012). Will disruptive mood dysregulation disorder reduce false diagnosis of bipolar disorder in children? Bipolar Disorders, 14(5), 488–496. doi:10.1111/j.1399-5618.2012.01029.x [PubMed: 22713098]

Mayes SD, Mathiowetz C, Kokotovich C, Waxmonsky J, Baweja R, Calhoun SL, & Bixler EO (2015). Stability of disruptive mood dysregulation disorder symptoms (irritable-angry mood and temper outbursts) throughout childhood and adolescence in a general population sample. Journal of Abnormal Child Psychology, 43(8), 1543–1549. doi:10.1007/s10802-015-0033-8 [PubMed: 26004122]

Mayes SD, Waxmonsky JD, Calhoun SL, & Bixler EO (2016). Disruptive mood dysregulation disorder symptoms and association with oppositional defiant and other disorders in a general population child sample. Journal of Child and Adolescent Psychopharmacology, 26(2), 101–106. doi:10.1089/cap.2015.0074 [PubMed: 26745442]

McGraw KO, & Wong SP (1996). Forming inferences about some intraclass correlation coefficients. Psychological Methods, 1(1), 30.

Mitchell RH, Timmins V, Collins J, Scavone A, Iskric A, & Goldstein BI (2016). Prevalence and correlates of disruptive mood dysregulation disorder among adolescents with bipolar disorder. Journal of Child and Adolescent Psychopharmacology, 26(2), 147–153. doi:10.1089/cap.2015.0063 [PubMed: 26844707]

PARS, Research Units on Pediatric Psychopharmacology Anxiety Study Group. (2002). The pediatric anxiety rating scale (PARS): Development and psychometric properties. Journal of the American Academy of Child & Adolescent Psychiatry, 41(9), 1061–1069. [PubMed: 12218427]

Poznanski EO, Grossman JA, Buchsbaum Y, Banegas M, Freeman L, & Gibbons R (1984). Preliminary studies of the reliability and validity of the children's depression rating scale. Journal of the American Academy of Child Psychiatry, 23(2), 191–197. [PubMed: 6715741]

Roy AK, Lopes V, & Klein RG (2014). Disruptive mood dysregulation disorder: A new diagnostic approach to chronic irritability in youth. American Journal of Psychiatry, 171(9), 918–924. doi:10.1176/appi.ajp.2014.13101301 [PubMed: 25178749]

Saylor CF, Finch AJ Jr., Spirito A, & Bennett B (1984). The children's depression inventory: A systematic evaluation of psychometric properties. Journal of Consulting and Clinical Psychology, 52(6), 955–967. doi:10.1037//0022-006x.52.6.955 [PubMed: 6520288]

Shaffer D, Gould MS, Brasic J, Ambrosini P, Fisher P, Bird H, & Aluwahlia S (1983). A children's global assessment scale (CGAS). Archives of General Psychiatry, 40(11), 1228–1231. doi:10.1001/archpsyc.1983.01790100074010 [PubMed: 6639293]

Shrout PE, & Fleiss JL (1979). Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin, 86(2), 420–428. [PubMed: 18839484]

Steiger JH (1980). Tests for comparing elements of a correlation matrix. Psychological Bulletin, 87, 245–251. doi:10.1037/0033-2909.87.2.245

Stoddard J, Sharif-Askary B, Harkins EA, Frank HR, Brotman MA, Penton-Voak IS, … Leibenluft E (2016). An Open Pilot Study of Training Hostile Interpretation Bias to Treat Disruptive Mood Dysregulation Disorder. Journal of Child and Adolescent Psychopharmacology, 26(1), 49–57. doi:10.1089/cap.2015.0100 [PubMed: 26745832]

Stoddard J, Stringaris A, Brotman MA, Montville D, Pine DS, & Leibenluft E (2014). Irritability in child and adolescent anxiety disorders. Depression and Anxiety, 31(7), 566–573. doi:10.1002/da.22151 [PubMed: 23818321]

Stringaris A (2011). Irritability in children and adolescents: A challenge for DSM-5. European Child and Adolescent Psychiatry, 20(2), 61–66. doi:10.1007/s00787-010-0150-4 [PubMed: 21298306]

Stringaris A (2013). Editorial: The new DSM is coming—it needs tough love. Journal of Child Psychology and Psychiatry and Allied Disciplines, 54(5), 501–502. doi:10.1111/jcpp.12078

Stringaris A, Cohen P, Pine DS, & Leibenluft E (2009). Adult outcomes of youth irritability: A 20-year prospective community-based study. American Journal of Psychiatry, 166(9), 1048–1054. doi:10.1176/appi.ajp.2009.08121849 [PubMed: 19570932]

Stringaris A, Goodman R, Ferdinando S, Razdan V, Muhrer E, Leibenluft E, & Brotman MA (2012). The Affective Reactivity Index: A concise irritability scale for clinical and research settings. Journal of Child Psychology and Psychiatry and Allied Disciplines, 53(11), 1109–1117. doi:10.1111/j.1469-7610.2012.02561.x

Stringaris A, Maughan B, Copeland WS, Costello EJ, & Angold A (2013). Irritable mood as a symptom of depression in youth: Prevalence, developmental, and clinical correlates in the Great Smoky Mountains Study. Journal of the American Academy of Child & Adolescent Psychiatry, 52(8), 831–840. doi:10.1016/j.jaac.2013.05.017 [PubMed: 23880493]

Towbin K, Vidal-Ribas P, Brotman MA, Pickles A, Miller KV, Kaiser A, … Stringaris A (2019). A double-blind randomized placebo-controlled trial of citalopram adjunctive to stimulant medication in youth with chronic severe irritability. Journal of the American Academy of Child & Adolescent Psychiatry. doi:10.1016/j.jaac.2019.05.015

Tufan E, Topal Z, Demir N, Taskiran S, Savci U, Cansiz MA, & Semerci B (2016). Sociodemographic and clinical features of disruptive mood dysregulation disorder: A chart review. Journal of Child and Adolescent Psychopharmacology, 26(2), 94–100. doi:10.1089/cap.2015.0004 [PubMed: 26491995]

Vidal-Ribas P, Brotman MA, Valdivieso I, Leibenluft E, & Stringaris A (2016). The status of irritability in psychiatry: A conceptual and quantitative review. Journal of the American Academy of Child & Adolescent Psychiatry, 55(7), 556–570. doi:10.1016/j.jaac.2016.04.014 [PubMed: 27343883]

Walkup JT, Labellarte MJ, Riddle MA, Pine DS, Greenhill L, Klein R, … Hack S. J. N. E. J. o. M. (2001). Fluvoxamine for the treatment of anxiety disorders in children and adolescents. New England Journal of Medicine, 344(17), 1279–1285. [PubMed: 11323729]

Waxmonsky JG, Wymbs FA, Pariseau ME, Belin PJ, Waschbusch DA, Babocsai L, … Pelham WE (2013). A novel group therapy for children with ADHD and severe mood dysregulation. Journal of Attention Disorders, 17(6), 527–541. doi:10.1177/1087054711433423 [PubMed: 22373865]

Wechsler D (1999). Manual for the Wechsler abbreviated intelligence scale (WASI). San Antonio, TX: The Psychological Corporation.

Wiggins JL, Brotman MA, Adleman NE, Kim P, Oakes AH, Reynolds RC, … Leibenluft E (2016). Neural correlates of irritability in disruptive mood dysregulation and bipolar disorders. American Journal of Psychiatry, 173(7), 722–730. doi:10.1176/appi.ajp.2015.15060833 [PubMed: 26892942]

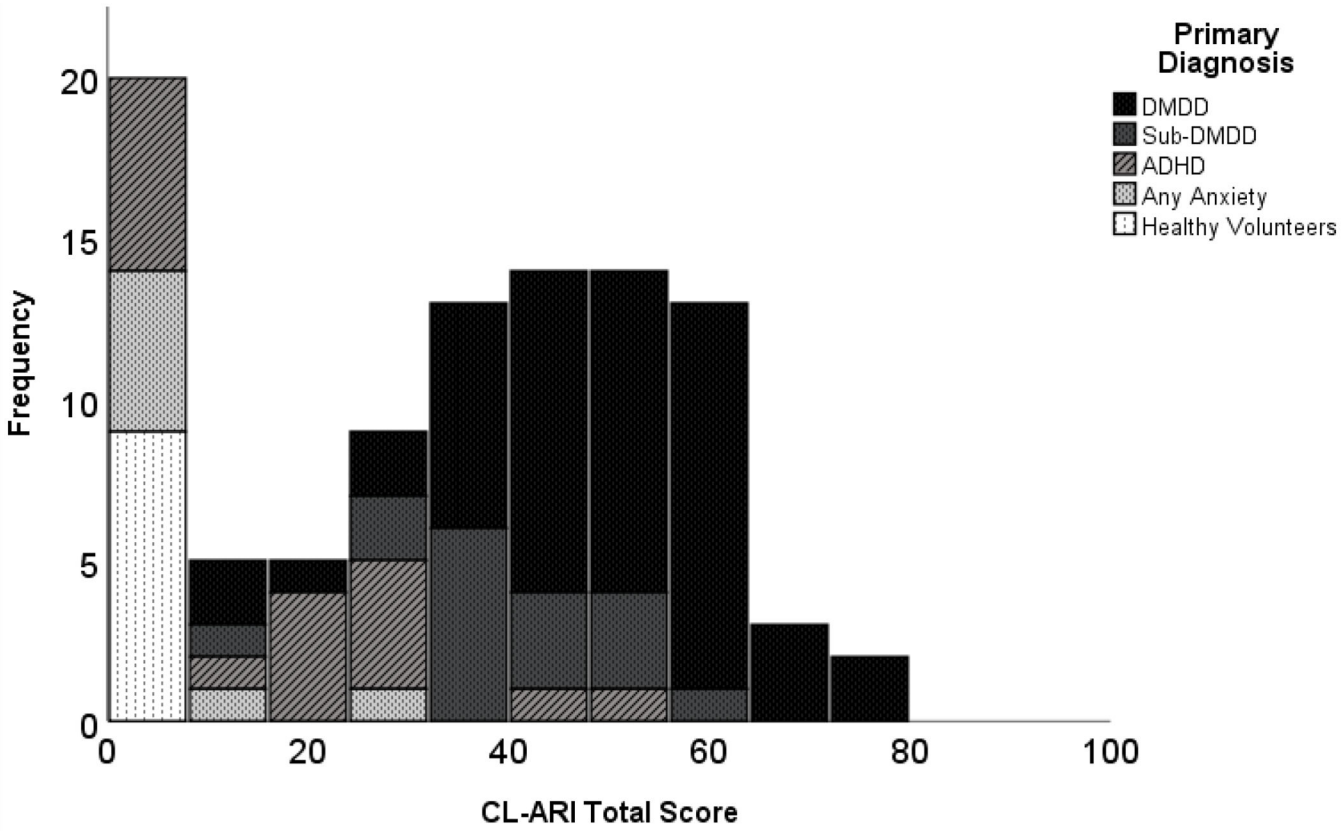

Figure 1.
CL-ARI total score by primary presenting diagnosis.

Table 1

Demographic and Clinical Characteristics of the Sample ($N = 98$)

Demographics and Clinical Characteristics	
N	98
Age *M* (SD)	12.66 (2.47)
Sex (% female)	41
Race (%)	
Asian	1
Black or African American	10.20
White	76.50
Other	12.30
Ethnicity (%)	
Latino or Hispanic	11.20
IQ *M* (SD)	110.91 (13.47)
Primary presenting diagnoses	
Any Anxiety Disorder (%)	7.10
ADHD	17.30
DMDD	50
subthreshold DMDD	16.30
No diagnosis	9.20
Lifetime Diagnoses (based on K-SADS; %)	
DMDD only	10.20
Any Anxiety Disorder only	6.10
ADHD only	18.40
DMDD and Any Anxiety Disorder only	7.10
DMDD and ADHD only	11.20
Any Anxiety Disorder and ADHD	6.10
DMDD, Any Anxiety Disorder and ADHD	24.50
No diagnosis	16.30

Note. ADHD: Attention deficit/hyperactivity disorder, DMDD: Disruptive mood dysregulation disorder

Table 2

CL-ARI Total Subscale, and Item Scores

CL-ARI Item (range, *M*, *SD*)	DMDD (*n* = 49)	Full Sample (*n* = 98)
Temper outbursts[a]		
Mild Frequency (0-4)	3.35 (1.17)	2.90 (1.47)
Mild Duration (0-5)	1.65 (0.86)	1.35 (0.86)
Moderate Frequency (0-4)	2.04 (0.98)	1.46 (1.19)
Moderate Duration (0-5)	2.12 (1.11)	1.51 (1.25)
Severe Frequency (0-4)	0.78 (0.90)	0.50 (0.83)
Severe Duration (0-5)	1.27 (1.52)	0.72 (1.24)
Total temper outburst subscale	11.20 (3.42)	8.44 (4.80)
Irritable mood		
Irritable mood Frequency (0-3)	2.61 (0.91)	1.87 (1.35)
Irritable mood Severity (0-5)	2.45 (1.00)	1.79 (1.35)
Total irritable mood subscale	5.06 (1.65)	3.65 (2.50)
Impairment		
Family (0-5)	3.27 (1.11)	2.24 (1.57)
School(0-5)	1.47 (1.71)	0.84 (1.46)
Peers(0-5)	1.59 (1.63)	1.00 (1.41)
Total impairment subscale	6.33 (3.36)	4.08 (3.62)
Total score	48.98 (14.32)	34.71 (21.85)

Note. DMDD: Disruptive mood dysregulation disorder

[a]For cases in which no outbursts were reported of a specific severity, duration for that severity was rated as 0.

Temper outbursts frequency scale: 0 = None, 1 = 1-2 outbursts this week, 2 = 3 or more outbursts this week, 3 = outbursts every day this week, 4 = more than 1 outburst every day this week.

Temper outbursts duration scale: 0 = None, 1 = <5 min, 2 = 5-10 mins, 3 = 11-30 mins, 4 = 31-60 mins, 5>60 mins.

Mood frequency scale: 0 = None, 1 = 1 day, 2 = 2-3 days. 3 = 4 or more days.

Mood severity scale: 0 Not present, 1 = Mild – Occasional, brief periods of slightly irritable mood or easily annoyed out of proportion to circumstances, 2 = Mild – Moderate, 3 = Moderate - Some sustained periods of irritable mood or angry affect, 4 = Moderate – Severe, 5 = Severe - Intense, almost constant feelings of anger or persistently irritable mood.

Table 3

Descriptive Statistics and Correlation Coefficients Between the CL-ARI and Other Clinician and Youth-Report Measures of Psychopathology and Impairment

Measure	Focus	Rater	*M* (SD)	*n*	*r*
CGI-S	Irritability	Clinician	3.11 (1.54)	90	.89***[a]
CGAS	Impairment	Clinician	57.08 (14.39)	90	−.75***[a]
PARS	Anxiety	Clinician	9.14 (6.06)	90	.51***[a]
CDRS	Depression	Clinician	25.33 (6.95)	90	.52***[a]
ARI-P	Irritability	Parent	7.54 (4.2)	71	.68***[a,b]
ARI-C	Irritability	Youth	5.27 (3.60)	71	.42***[a,b]
SCARED-P	Anxiety	Parent	16.75 (12.60)	71	.32**[a]
SCARED-C	Anxiety	Youth	18.92 (11.40)	71	.09
CDI	Depression	Youth	10.11 (7.36)	45	.39**[a]

Note. CGI-S: Clinical Global Impressions Scale-Severity (DMDD), CGAS: Children's Global Assessment Scale, PARS: Pediatric Anxiety Rating Scale, CDRS: Children's Depression Rating Scale, ARI-P: Affective Reactivity Index-Parent, ARI-C: Affective Reactivity Index-Child, SCARED-P: Screen for Child Anxiety Related Emotional Disorders-Parent, SCARED-C: Screen for Child Anxiety Related Emotional Disorders-Child, CDI: Child Depression Inventory

* $p<.05$,

** $p<.01$,

*** $p<.001$

[a] Correlations remain statistically significant when controlling for age in a partial correlation.

[b] ARI-P and ARI-C scoring procedures usually do not include the impairment item in total scores, hence, additional analyses without impairment items were performed: CL-ARI and ARI-P: $r(69) = .69$, $p<.001$, CL-ARI and ARI-C: $r(69) = .48$, $p<.001$