

RESEARCH

Open Access



Assessment of affective dysregulation in children: development and evaluation of a semi-structured interview for parents and for children

Anne-Katrin Treier^{1,2*}, Sara Zaplana Labarga^{1,2}, Claudia Ginsberg^{1,2}, Lea Teresa Kohl^{1,2}, Anja Görtz-Dorten^{1,2}, Ulrike Ravens-Sieberer³, Anne Kaman³, Tobias Banaschewski⁴, Pascal-M. Aggensteiner⁴, Charlotte Hanisch⁵, Michael Kölch^{6,7}, Andrea Daunke⁷, Veit Roessner⁸, Gregor Kohls⁸ on behalf of the ADOPT consortium, and Manfred Döpfner^{1,2}

Abstract

Background Children with affective dysregulation (AD) show an excessive reactivity to emotionally positive or negative stimuli, typically manifesting in chronic irritability, severe temper tantrums, and sudden mood swings. AD shows a large overlap with externalizing and internalizing disorders. Given its transdiagnostic nature, AD cannot be reliably and validly captured only by diagnostic categories such as disruptive mood dysregulation disorder (DMDD). Therefore, this study aimed to evaluate two semi-structured clinical interviews—one for parents and one for children.

Methods Both interviews were developed based on existing measures that capture particular aspects of AD. We analyzed internal consistencies and interrater agreement to evaluate their reliability. Furthermore, we analyzed factor loadings in an exploratory factor analysis, differences in interview scores between children with and without co-occurring internalizing and externalizing disorders, and associations with other measures of AD and of AD-related constructs. The evaluation was performed in a screened community sample of children aged 8–12 years ($n = 445$). Interrater reliability was additionally analyzed in an outpatient sample of children aged 8–12 years ($n = 27$).

Results Overall, internal consistency was acceptable to good. In both samples, we found moderate to excellent interrater reliability on a dimensional level. Interrater agreement for the dichotomous diagnosis DMDD was substantial to perfect. In the exploratory factor analysis, almost all factor loadings were acceptable. Children with a diagnosis of disruptive disorder, attention-deficit/hyperactivity disorder, or any disorder (disruptive disorder, attention-

The members of the ADOPT consortium group are listed in the Acknowledgement section.

*Correspondence:
Anne-Katrin Treier
anne-katrin.treier@uk-koeln.de

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

deficit/hyperactivity disorder, and depressive disorder) showed higher scores on the DADYS interviews than children without these disorders. The correlation analyses revealed the strongest associations with other measures of AD and measures of AD-specific functional impairment. Moreover, we found moderate to very large associations with internalizing and externalizing symptoms and moderate to large associations with emotion regulation strategies and health-related quality of life.

Conclusions The analyses of internal consistency and interrater agreement support the reliability of both clinical interviews. Furthermore, exploratory factor analysis, discriminant analyses, and correlation analyses support the interviews' factorial, discriminant, concurrent, convergent, and divergent validity. The interviews might thus contribute to the reliable and valid identification of children with AD and the assessment of treatment responses.

Trial registration ADOPT Online: German Clinical Trials Register (DRKS) DRKS00014963. Registered 27 June 2018.

Keywords Affective dysregulation, Irritability, Children, Assessment, Clinical interview, Reliability, Validity

Introduction

Children with affective dysregulation (AD) typically show chronic irritability, severe temper tantrums, and sudden mood swings [1–3]. Emotion recognition and regulation develop from birth through interaction with a sensitive caregiver and lead to primary regulation strategies at the age of seven, which become more self-directed with increasing age [4]. Dysfunctions of emotion recognition and/or emotion regulation are suggested underlying mechanisms of AD, an assumption that is supported by findings of an elevated use of maladaptive emotion regulation strategies in children with AD [5]. In contrast to the concept of irritability—which solely comprises the proneness to anger [6]—AD additionally encompasses emotional reactivity such as anxiety, sadness, or positive emotions (e.g., exuberance; [3]).

However, there are various different operationalizations of AD. The fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; [7]) introduced the diagnosis of disruptive mood dysregulation disorder (DMDD) as a categorical diagnosis for children with irritability and severe temper tantrums. In community samples, between 0.8% and 9% of all children and adolescents fulfill the diagnostic criteria for DMDD [8, 9], with lower rates in clinical ratings and higher rates in parent ratings. Furthermore, symptoms of AD in early childhood can be categorized under the disorder of dysregulated anger and aggression of early childhood based on the second revision of the Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood (DC:0–5; [10]) from the age of 24 months onwards [11]. In the 11th revision of the International Classification of Diseases (ICD-11; [12]), AD is considered a diagnostic specifier of oppositional defiant disorder (ODD). According to the ICD-11, ODD can be defined with and without chronic irritability, due to the large overlap of diagnostic criteria between ODD and AD. Nevertheless, AD symptoms do not only occur in patients with ODD. On the contrary, children and adolescents with AD often show other externalizing

disorders, especially attention-deficit/hyperactivity disorder (ADHD) but also conduct disorder (CD), as well as internalizing disorders, especially depressive and anxiety disorders [3, 8, 13].

Due to the overlap with this broad spectrum of other disorders, AD is also conceptualized as a transdiagnostic and dimensional rather than a distinct, categorical phenomenon [2, 14]. While the categorical classification of a disorder is useful in terms of guiding empirical research and decision-making such as treatment indication and selection, for the substantial number of patients with co-occurring disorders, such decisions might be more complicated [15]. In view of the large overlap of AD with both externalizing and internalizing disorders, a transdiagnostic and dimensional approach might therefore be more appropriate [14]. A stronger dimensional approach is also supported by the Research Domain Criteria (RDoC) initiative of the National Institute of Mental Health [16]. Within this initiative, AD fits well in the concept of frustrative non-reward in the negative emotionality domain [17].

In summary, we adopt a broader, dimensional, and transdiagnostic concept of AD, and perceive AD as excessive reactivity to emotionally positive or negative stimuli [2, 3]. Accordingly, AD comprises a proneness to a variety of emotional reactions, ranging from anger to anxiety or sadness, but also includes positive emotions such as exuberance.

Since the concept of AD is rather new, currently available instruments only assess particular aspects of AD: There are several questionnaires, which focus on irritability (Affective Reactivity Index, ARI; [18]), anger (Patient-Reported Outcome Measurement Information System Anger Scale, PROMIS; [19]), emotion regulation (Emotion Regulation Checklist, ERC; [20]), or DMDD (Diagnostic System for Mental Disorders in children and adolescents according to ICD-10 and DSM-5, DISYPS-III; [21]) Additionally, to assess the so-called dysregulation profile, there are two broadband questionnaires: the Child Behavior Checklist–Dysregulation

Profile (CBCL-DP; [22]) and the Strengths and Difficulties Questionnaire–Dysregulation Profile (SDQ-DP; [23, 24]), with the dysregulation profile being defined as the co-existence of anxious/depressive, attention, and aggressive problems [22]. In both of these questionnaires, the profile is formed by combining specific subscales/items [22–24]. Finally, there are several clinical interviews, which focus on irritability (Clinician Affective Reactivity Index, CL-ARI; [25]) and DMDD (DMDD module of the Kiddie Schedule for Affective Disorders and Schizophrenia for School Aged Children Present and Lifetime Version, K-SADS-PL; [26], DMDD module of the Extended Strengths and Weaknesses Assessment of Normal Behavior, E-SWAN; [27]). Overall, however, comprehensive analyses of both the reliability and validity of instruments assessing AD are lacking (see the systematic review of measures assessing DMDD by [28]).

To date, the only comprehensive tool focusing on the broad conceptualization of AD is the German-language Diagnostic Tool for Affective Dysregulation in Children (DADYS; [3, 14, 29, 30]). The DADYS covers symptoms of irritability, impulsivity, temper outbursts, anger, mood swings, sadness, and exuberance. Besides a screening questionnaire (DADYS-Screen; [3, 30]), the DADYS comprises a parent, a teacher, and a child questionnaire (DADYS-PQ, DADYS-TQ, DADYS-CQ; [14, 29]), as well as semi-structured parent and child interviews (DADYS-PI, DADYS-CI; [29]). While the DADYS-Screen and the DADYS-PQ have already been comprehensively evaluated [3, 14, 30], the reliability and validity of the DADYS interviews have not yet been assessed.

Accordingly, the present study aimed to evaluate the DADYS interviews. The study was conducted in the context of the multicenter study ADOPT (Affective Dysregulation–Optimizing Prevention and Treatment), which integrates internationally established, highly experienced, and interdisciplinary research groups [2]. First, we analyzed the factor structure of the interviews by means of exploratory factor analyses. Second, we analyzed internal consistency and interrater reliability of the scales developed on the basis of factor analyses. Third, we aimed at demonstrating the validity of the DADYS interviews by analyzing their discriminative power to differentiate categorically between children with and without co-occurring internalizing and externalizing disorders (discriminant validity) as well as the dimensional associations with other measures of AD (concurrent validity) and with measures assessing emotion regulation strategies, externalizing and internalizing symptoms, and health-related quality of life (convergent and divergent validity).

Methods

Participants

For the evaluation of the DADYS interviews, two samples were used: A *screened community sample* ($n=445$) was formed from a larger community sample ($n=9,759$) which was recruited through residents' registration offices in four German cities. Children aged 8 to 12 years were screened using the parent-rated DADYS-Screen [3, 30]. The age range of 8 to 12 years was chosen due to the ADOPT study's focus on AD in childhood and because we wanted to administer the DADYS in both the clinical child interview and the child questionnaire. After screening for AD, children were classified as high AD (highest 10% raw scores; AD) or low AD (lowest 10% raw scores, NoAD). We chose this cut-off of 10% as an approximation of the prevalence rates of up to 9% reported in epidemiological studies [9]. The families of all children with AD and a random sample of the families of children without AD were invited to participate in a comprehensive assessment and were screened for additional inclusion criteria in one of the five study centers (child and adolescent psychiatric units or outpatient units). All families with AD fulfilling the additional criteria were invited to participate in the subsequent treatment study receiving either an AD-specific treatment or treatment as usual, while families without AD were subsequently monitored as a comparison group (ADOPT study, [2]). Additional inclusion/exclusion criteria were no participation in behavioral therapy focusing on AD, no autism spectrum disorder, and an intelligence quotient (IQ, based on clinical judgment) above 80. The assessment was completed online using the REDCap electronic data capture tool or offline in paper-and-pencil format. If parents provided permission, the DADYS interview was audio- or videotaped. For the assessment of interrater reliability, audio- or videotaped parent interviews ($n=246$) were additionally rated by a blinded rater who was blind to the group status and the time of measurement.

An *outpatient sample* ($n=27$) was recruited in the outpatient unit of the study center in Cologne for the analysis of interrater reliability of both the parent and the child interview in a clinical sample. All participating children had at least one diagnosis according to DSM-5. If parents and children provided permission, the DADYS interview was audio- or videotaped ($n=27$). Participating families were at different stages of their psychotherapeutic treatment.

The unblinded interviewers at all study centers had a Bachelor's or Master's degree in psychology or education. All unblinded interviewers received standardized and extensive training on conducting the DADYS interviews, including practice videos for each item, sitting in on interviews, and conducting interviews under supervision. Blinded ratings and all ratings of the outpatient

sample were conducted at both study sites in Cologne. Blinded raters for interrater agreement received the same extensive training, with the exception that they did not sit in on interviews or conduct interviews themselves. All unblinded interviewers and blinded raters were encouraged to consult their supervisor if they experienced any difficulties regarding the assessment with the DADYS interviews.

DADYS parent and child interviews

The DADYS parent (DADYS-PI) and the parallel child interview (DADYS-CI) are semi-structured interviews which aim to calculate a dimensional score for AD. The DADYS-PI additionally allows for the categorical assessment of a DMDD diagnosis. To operationalize the broad conceptualization, we developed an item pool based on existing measures with different foci (ARI, [18], ERC, [20], DISYPS-III, [21], Conners' Rating Scale, [31]). Items of the DISYPS-III and the ARI were combined due to item overlap. The original item pool comprised 35 items. For item selection, a Delphi rating of experts, and focus groups with experts and parents were implemented [30]. In the first step, items with very similar content, as rated by clinical experts in the field of AD (MD, AGD, URS), were deleted. In the second step, items were further reduced by conducting a focus group with clinical experts (clinicians and psychotherapists), a focus group with parents (outpatient clinic), and a focus group with children (outpatient clinic). The item set used in the sample for evaluating the interview consisted of 13 items assessing symptoms and five items assessing functional impairment for the DADYS-PI (see Table S1 in the supplement) as well as 10 items assessing symptoms and five items assessing functional impairment for the DADYS-CI.

In the two interviews, parents and children were asked to describe in detail the respective emotional or behavioral response defined in each item (e.g., exhibits strong mood swings, is able to delay gratification, is overly exuberant, is quick to anger) and to describe specific situations in which the response might be observed as well as the frequency of the response. Each item was rated by the clinician on a 4-point Likert scale ranging from 0 (age-appropriate/not present) to 3 (very strongly present). For each score, a brief description of the symptom severity was provided to aid clinical judgment. Items on functional impairment were only assessed if AD symptoms were present. For the assessment of the DMDD diagnosis in the DADYS-PI, the age of symptom onset, duration of symptoms, pervasiveness, and exclusion criteria were additionally assessed. Lastly, the severity of AD was rated globally with one item in the DADYS-PI.

Measures for the validity assessment

Affective dysregulation. For the assessment of AD, two measures were used in addition to the DADYS interviews: (a) the DADYS parent questionnaire (DADYS-PQ) and child questionnaire (DADYS-CQ; [14, 29]) as well as (b) the Dysregulation Profile of the Child Behavior Checklist in its German version (CBCL-DP; [22, 32]). **DADYS-PQ/-CQ.** The DADYS questionnaires were also developed based on existing measures assessing aspects of AD (ARI, [18], PROMIS, [19], ERC, [20], DISYPS-III, [21], Conners' Rating Scale, [31]), with an overlap with the original item pool of the DADYS interview. Items for the DADYS-Screen were selected and evaluated using a mixed-methods approach, including a Delphi rating of experts, focus groups with parents and experts, and psychometric analyses. Each item was rated on a 4-point Likert scale ranging from 0 (age-appropriate/not present) to 3 (very strongly present). For each questionnaire, the mean item score was calculated for the total symptoms scale (DADYS-PQ: 27 items; DADYS-CQ: 26 items) and for the functional impairment scale (DADYS-PQ: 5 items; DADYS-CQ: 5 items). In the current screened community sample, the internal consistency of each scale was sufficient to excellent ($0.77 \leq \alpha \leq 0.94$). **CBCL-DP.** Parents rated the items of the subscales attention problems (10 items), aggressive behavior (18 items), and anxious/depressed (13 items) on a 3-point Likert scale ranging from 0 (not true) to 2 (very true or often true). We calculated the CBCL-DP scale by summing the mean item score for each subscale (range 0–2), which resulted in a range from 0 to 6 [33]. In the current screened community sample, the CBCL-DP scale showed excellent internal consistency ($\alpha = 0.94$).

Emotion regulation. The Questionnaire for the Regulation of Frustration in Children was used to assess emotion regulation strategies (FRUST; [34, 35]). The questionnaire comprises the two subscales adaptive and maladaptive emotion regulation strategies, and is rated by parent (adaptive strategies: 10 items, maladaptive strategies: 4 items) and child (adaptive strategies: 33 items, maladaptive strategies: 7 items). The adaptive subscale includes strategies such as problem-solving or social support and the maladaptive subscale includes strategies such as rumination or avoidance. The items are rated on a 5-point Likert scale ranging from 0 (hardly ever) to 4 (almost always). We calculated a mean item score for each subscale. The internal consistency of the two subscales in the current screened community sample was sufficient to excellent ($0.78 \leq \alpha \leq 0.94$).

Externalizing/internalizing symptoms. To assess the children's internalizing and externalizing symptoms, we used the DISYPS-III [21]. Specifically, we employed the therapist-rated diagnostic screening checklists for internalizing symptoms (ILF-SCREEN-Internal, 19 items)

and externalizing symptoms (ILF-SCREEN-External, 9 items), based on a parent interview. Additionally, we used the parent- and child-rated symptom checklists for ADHD (20 items) and disruptive disorders—including ODD, CD, DMDD, and callous-unemotional traits—(28 items). All items were rated on a 4-point Likert scale ranging from 0 (age-typical) to 3 (very strong). A mean item score across all items was calculated for each checklist, with the exception of the checklist for disruptive disorders, for which we calculated subscale scores for ODD (8 items), and CD (6 items). Since the items of the DMDD scale were part of the DADYS-PQ, we did not calculate this scale separately for the present study. In the current screened community sample, internal consistency was good to excellent for clinician-rated internalizing and externalizing symptoms as well as for parent- and child-rated ADHD and ODD symptoms ($0.82 \leq \alpha \leq 0.95$), except for the parent- and child-rated CD scale ($0.59 \leq \alpha \leq 0.60$) due to the diverse range of behaviors assessed in this scale.

Furthermore, due to the especially frequent co-occurrence of AD with both externalizing and affective disorders, we assessed diagnoses of ADHD, disruptive disorders (ODD and/or CD), and depressive disorder, coded as 0 (no) and 1 (yes), based on the DISYPS parent interviews. If the parents reported symptoms of these disorders on the screening checklist, we verified the respective diagnosis using the comprehensive checklist from the DISYPS-III.

Quality of life. KIDSCREEN. To assess health-related quality of life, we used the KIDSCREEN questionnaire [58], which measures subjective health and well-being in children and adolescents. Children completed the KIDSCREEN-10 Index (10 items) and parents completed the short version KIDSCREEN-27 (27 items). For both the child and parent version, items are rated on a 5-point Likert scale ranging from 1 (never/not at all) to 5 (always/very strong). The mean item score was calculated. Internal consistency in the current screened community sample was good to excellent ($0.81 \leq \alpha \leq 0.91$).

Statistical analyses

All statistical analyses were performed using SPSS Version 29 [36]. To reduce bias in the results, scales were only computed if at least 90% of the respective scale items were available [37].

Sample characteristics. Differences in sample characteristics between the AD and the NoAD subsample were examined using χ^2 tests for categorical variables, t -tests for interval-scaled variables, and Kruskal-Wallis tests for ordinal-scaled variables. As measures of effect size, we used Cramer's V for χ^2 tests ($0.10 \leq \phi_c < 0.30$ small, $0.30 \leq \phi_c < 0.50$ moderate, $0.50 \leq \phi_c$ large), Cohen's d [38] for t -tests ($0.20 \leq d < 0.50$ small, $0.50 \leq d < 0.80$ moderate,

$0.80 \leq d$ large), and Pearson correlations for Kruskal-Wallis tests ($0.10 \leq r < 0.30$ small, $0.30 \leq r < 0.50$ moderate, $0.50 \leq r$ large; [38]).

Exploratory factor analysis. For item reduction and scale development as well as the analysis of factor loadings (factorial validity), we performed an exploratory factor analysis for both the DADYS-PI and the DADYS-CI in the screened community sample. For each interview, we performed a principal component analysis (PCA) and a principal factor analysis (PFA). Only symptom items were included in the factor analysis since the functional impairment scale of the DISYPS-III was derived as a whole [21]. The Kaiser-Meyer-Olkin measure of sampling adequacy resulted in superb values ($KMO_{DADYS-PI} = 0.94$; $KMO_{DADYS-CI} = 0.92$). The scree test [39], the MAP test [40], and the parallel analysis [41] were used to determine the number of factors. Factor loadings of $a \geq 0.30$ in the PCA and the PFA were considered robust [42].

Scale characteristics. To evaluate internal consistency, Cronbach's alpha was calculated for the total symptom scale and the functional impairment scale for both interviews in the screened community sample, with values of $a > 0.70$ considered acceptable [43]. Furthermore, the corrected item-total correlation was calculated for each item, with values of $r > 0.30$ considered acceptable [44].

Interrater reliability. Interrater reliability for continuous AD symptoms was evaluated using intraclass correlations (ICC; [45]). Since the characteristics of the data differed between the screened community sample and the outpatient sample, we applied different models to calculate the ICC in the two samples. To compare unblinded interviewer ratings and blinded ratings in the screened community sample, we calculated the ICC one-way random-effects, absolute agreement model for single-rater ICC(1,1), as the multicenter design of the study did not allow for the same unblinded interviewers for all patients [46]. To compare the ratings in the outpatient sample, we calculated the ICC two-way random effects, absolute agreement, and single-rater ICC(2,1), as we used the same group of raters for all patients [46]. Single-rater models were applied in both samples as they more appropriately reflect routine clinical care, where one clinician usually conducts ratings [46]. For the interpretation of ICC values, we followed Koo and Li (ICC < 0.50 poor, $0.50 \leq ICC \leq 0.74$ moderate, $0.75 \leq ICC \leq 0.89$ good, $ICC > 0.90$ excellent; [46]). Interrater reliability for the dichotomous DMDD diagnosis based on the DADYS-PI was evaluated using Cohen's kappa [47] in both samples. For the interpretation of ICC values, we followed Landis and Koch ($\kappa < 0.20$ slight, $0.21 \leq \kappa \leq 0.40$ fair, $0.41 \leq \kappa \leq 0.60$ moderate, $0.61 \leq \kappa \leq 0.80$ substantial, $\kappa > 0.80$ almost perfect; [48]).

Discriminant validity. To analyze discriminant validity, we evaluated the differences between the DADYS

interview scores of children with and without a diagnosis of disruptive disorders (ODD or CD), ADHD, depressive disorder, or an overarching diagnosis of any of these disorders using *t*-tests. Cohen’s *d* [38] was applied as a measure of effect size, using the interpretation mentioned above.

Concurrent, convergent, and divergent Validity. To analyze concurrent, convergent, and divergent validity, we calculated Pearson correlations between the DADYS interviews and comprehensive, parent- and child-rated measures of AD, parent- and child-rated measures of emotion regulation strategies, parent-, child-, and clinician-rated measures of externalizing and internalizing symptoms, and parent- and child-rated measures of health-related quality of life. Pearson correlation coefficients were interpreted as outlined above. Additionally, correlations accounting for at least 50% of the variance ($r > .70$) were classified as very large. Furthermore, we calculated paired *t*-tests for the comparison of mean differences between the two DADYS interviews and between the DADYS interviews and the DADYS questionnaire. Cohen’s *d* [38] was applied as a measure of effect size, with the aforementioned interpretation.

Results

Sample characteristics

The total screened community sample had a mean age of 10.70 years ($SD=1.32$) and a mean socioeconomic status of 6.26 ($SD=1.20$; range 1–7; value is based on the average national income obtained with the highest education and occupational qualification in the family; [49]) and 55.3% were boys. Approximately 32% of the children in the total screened community sample were diagnosed with ODD, CD, ADHD, or depressive disorder (see Table 1). When comparing the AD with the NoAD subsample of the screened community sample, we found a higher percentage of boys (small effect), lower age (small effect), and more co-occurring disorders (large effect) in the AD sample. The outpatient sample had a mean age of 10.41 years ($SD=1.45$) and a mean socioeconomic status of 4.83 ($SD=1.76$; range 1–7) and 81.5% were boys.

Almost all children in the outpatient sample (90.9%) were diagnosed with ODD, CD, ADHD, or depressive disorder.

Exploratory factor analysis

The scree test [39], the MAP test [40], and the parallel analysis [41] pointed to a one-factor solution for both interviews. Therefore, we specified the number of factors to one factor, and all symptom items were combined in the total (AD) symptom scale. The AD factor explained 44.94% of the variance in the DADYS-PI and 47.53% of the variance in the DADYS-CI.

For the DADYS-PI, factor loadings ranged from 0.39 to 0.87 ($M=0.65$, $SD=0.17$) in the PCA and from 0.35 to 0.88 ($M=0.62$, $SD=0.18$) in the PFA (see Table 2). For the DADYS-CI, factor loadings ranged from 0.30 to 0.86 ($M=0.67$, $SD=0.18$) in the PCA and from 0.25 to 0.85 ($M=0.63$, $SD=0.20$) in the PFA. The lowest factor loadings in each analysis were found for the item “exuberance”. Since this item fell below our predefined robustness criterion of $\alpha=0.30$ [42] in only one of four analyses, and as we considered it important for the broader concept of AD, we decided to retain this item. All other items were considered robust and likewise retained.

Scale characteristics

All items of the DADYS-PI and the DADYS-CI demonstrated the full scale range from 0 to 3. Item mean scores on the total symptom scale ranged from 0.33 (item “exuberance”) to 1.23 (item “quick to anger”) on the DADYS-PI and from 0.21 (item “exuberance”) to 1.00 (item “self-regulation”) on the DADYS-CI (see Table 2). Item mean scores on the functional impairment scale ranged from 0.45 (item “impaired relationships with adults”) to 1.79 (item “impaired relationships with family members”) on the DADYS-PI and from 0.19 (item “impaired academic performance”) to 0.81 (item “strain”) on the DADYS-CI.

Internal consistency for the total symptom scale was good in both interviews (DADYS-PI: $\alpha=0.89$; DADYS-CI: $\alpha=0.87$), whereas internal consistency for the functional impairment scale was sufficient in the DADYS-CI

Table 1 Participant characteristics for the screened community sample including children with (AD) and without affective dysregulation (No-AD) and the outpatient sample

	Range	Screened community sample			Test statistic	Effect size	Outpatient sample AD
		Total (n=445)	AD (n=265)	No-AD (n=180)			
Gender (male): %		55.3%	62.6%	44.4%	$\chi^2(1) = 14.36, p < .001$	$\phi_c = 0.18$	81.5%
Age (years): <i>M</i> (<i>SD</i>)	8–12	10.70 (1.32)	10.53 (1.28)	10.95 (1.34)	$t(443) = 3.33, p = .001$	$d = 0.32$	10.41 (1.45)
Socioeconomic status ^a : <i>M</i> (<i>SD</i>)	1–7	6.26 (1.20)	6.20 (1.23)	6.35 (1.17)	$U = 21266.50, p = .138$	-	4.83 (1.76)
Diagnosis ODD, CD, ADHD, or DD: %		32.4%	54.3%	0%	$\chi^2(1) = 144.61, p < .001$	$\phi_c = 0.57$	90.9%

Test statistics are based on χ^2 tests for categorical variables, *t*-tests for interval-scaled variables and Mann-Whitney U tests for ordinal-scaled variables. ϕ_c =effect size Cramer’s V for χ^2 tests. *d*=effect size Cohen’s *d* for *t*-tests. *M*=mean. *SD*=standard deviation. ODD=oppositional defiant disorder. CD=conduct disorder. ADHD=attention-deficit/hyperactivity disorder. DD=depressive disorder

^aValue is based on the average national income obtained with the highest education and occupational qualification in the family [49]

Table 2 Item statistics; interrater reliability, factor loadings, item-total correlation for DADYS parent interview (DADYS-P) and child interview (DADYS-CI)

Item	DADYS-PI					DADYS-CI							
	M ^a	SD ^a	ICC(1,1) ^b	ICC(2,1) ^c	PCA ^a	PFA ^a	Item-total ^a	M ^a	SD ^a	ICC(2,1) ^c	PCA ^a	PFA ^a	Item-total ^a
Mean total symptom scale	0.83	0.61	0.94	0.94	-	-	-	0.69	0.53	0.87	-	-	-
01 cheerfulness (recoded)	0.74	0.76	0.66	0.76	0.69	0.64	0.61	0.72	0.68	0.57	0.61	0.54	0.51
02 mood swings	1.22	1.11	0.85	0.87	0.87	0.87	0.81	0.75	0.82	0.87	0.83	0.81	0.74
03 self-regulation (recoded)	1.01	0.97	0.70	0.46	0.62	0.57	0.54	1.00	0.90	0.71	0.62	0.55	0.55
04 delay of gratification (recoded)	0.92	1.04	0.69	0.58	0.58	0.54	0.51	0.97	1.04	0.49	0.45	0.39	0.38
05 demanding	1.00	1.03	0.72	-	0.70	0.66	0.63	-	-	-	-	-	-
06 clinginess	0.49	0.74	0.80	0.76	0.47	0.42	0.40	-	-	-	-	-	-
07 exuberance	0.33	0.67	0.84	0.86	0.39	0.35	0.34	0.21	0.49	0.76	0.30	0.25	0.24
08 verbalizing emotions (recoded)	0.87	0.96	0.79	0.71	0.43	0.38	0.36	-	-	-	-	-	-
09 sadness/ listlessness	0.37	0.66	0.82	0.84	0.50	0.46	0.43	0.49	0.65	0.54	0.63	0.58	0.52
10 temper tantrums	1.09	1.01	0.87	0.64	0.85	0.85	0.80	0.80	0.89	0.45	0.85	0.85	0.77
11 quick to anger	1.23	1.11	0.90	0.88	0.87	0.88	0.82	0.81	0.88	0.58	0.86	0.85	0.77
12 offended	0.87	0.97	0.85	0.94	0.75	0.72	0.68	0.55	0.73	0.83	0.76	0.73	0.66
13 irritability	0.64	0.87	0.86	0.91	0.74	0.71	0.66	0.57	0.74	0.95	0.77	0.73	0.67
Mean functional impairment scale	1.05	0.52	0.63	0.85	-	-	-	0.45	0.48	0.72	-	-	-
14 impaired relationships with family members	1.79	0.81	0.57	0.82	-	-	0.22	0.73	0.85	0.91	-	-	0.56
15 impaired relationships with adults	0.50	0.76	0.60	0.68	-	-	0.38	0.20	0.51	0.80	-	-	0.39
16 impaired relationships with children; limited leisure activities	0.82	0.96	0.76	0.85	-	-	0.40	0.32	0.65	0.39	-	-	0.47
17 impaired academic performance	0.68	0.87	0.62	0.62	-	-	0.33	0.19	0.50	0.66	-	-	0.45
18 strain	1.45	0.87	0.59	0.76	-	-	0.33	0.81	0.93	0.81	-	-	0.58

DADYS=Diagnostic Tool for Affective Dysregulation in Children. M=mean, SD=standard deviation. ICC(1,1)=intraclass correlation coefficient with one-way random effects model, absolute agreement, single measurement. ICC(2,1)=intraclass correlation coefficient with two-way random effects model, absolute agreement, single measurement. PCA=principal component analysis (loadings). PFA=principal factor analysis (loadings). item-total=corrected item-total correlation

^aScreened community sample: DADYS-PI (total symptoms n=445, functional impairment n=262), DADYS-CI (total symptoms n=444, functional impairment n=376)

^bScreened community sample: blind DADYS-PI (total symptoms n=246, functional impairment n=139)

^cOutpatient sample: DADYS-PI (total symptoms n=25, functional impairment n=23), DADYS-CI (total symptoms n=21, functional impairment n=18)

($\alpha=0.72$) but insufficient in the DADYS-PI ($\alpha=0.57$). For the total symptom scale, item-total correlations were acceptable for all items in the DADYS-PI ($0.34 \leq r \leq .82$) and for all items in the DADYS-CI ($0.38 \leq r \leq .77$; see Table 2), with the exception of the item “exuberance” ($r = .24$). For the functional impairment scale, item-total correlations were acceptable for all items in the DADYS-PI ($0.33 \leq r \leq .40$), with the exception of the item “impaired relationships with family members” ($r = .22$), and for all items in the DADYS-CI ($0.39 \leq r \leq .58$).

Interrater reliability

The total symptom scale demonstrated good to excellent interrater reliability (screened community sample: DADYS-PI ICC[1,1]=0.94; outpatient sample: DADYS-PI ICC[2,1]=0.94, DADYS-CI ICC[2,1]=0.87; see Table 2). The functional impairment scale demonstrated moderate to good interrater reliability (screened community sample: DADYS-PI ICC[1,1]=0.63; outpatient sample: DADYS-PI ICC[2,1]=0.85, DADYS-CI ICC[2,1]=0.72).

We found a substantial interrater agreement for the DMDD diagnosis based on the DADYS-PI in the screened community sample ($\kappa=0.73$) and a perfect interrater agreement in the outpatient sample ($\kappa=1.00$).

Discriminant validity

Children with a disruptive disorder, ADHD, or any disorder (disruptive disorder, ADHD, or depressive disorder) scored higher than children without these disorders both on the DADYS-PI and the DADYS-CI (all large effects, see Table 3). As only 10 patients in the total sample showed a depressive disorder, we did not calculate the planned analyses for this disorder.

Concurrent, convergent, and divergent validity

The correlation between the DADYS-PI and the DADYS-CI was very large regarding the total symptom scales ($r = .77$) but moderate regarding the functional impairment scale ($r = .31$). Similarly, we found small mean differences between the DADYS-PI and the DADYS-CI for the total

symptom scale (full DADYS-PI scale: $d=0.36$; DADYS-PI cross-informant scale: $d=0.38$) and moderate differences for the functional impairment scale ($d=0.72$). Furthermore, the correlation between the total symptom scale and the functional impairment scale was moderate for the DADYS-PI ($r = .42$) and very large for the DADYS-CI ($r = .74$).

Associations between the DADYS interviews and other measures are presented in Table S2 in the Supplement. Regarding the association between the DADYS interviews (total symptom scale) and other measures of AD, we found large to very large correlations with parent- and child-rated DADYS questionnaires (total symptom scale) and with the parent-rated CBCL Dysregulation Profile (DADYS-PI: $0.64 \leq r < .87$; DADYS-CI: $0.64 \leq r < .79$; all $p < .01$). For the respective AD-specific functional impairment scale, we found small to large correlations between the DADYS interviews and the DADYS questionnaire (DADYS-PI: $0.26 \leq r < .54$; DADYS-CI: $0.49 \leq r < .67$; all $p < .01$). Furthermore, we found small to moderate mean differences between the DADYS interviews and the DADYS questionnaires for the total symptom scale (DADYS-PI/DADYS-PQ cross-informant scale: $d=0.48$; DADYS-CI/DADYS-CQ cross-informant scale: $d=0.68$), and small to no meaningful differences for the functional impairment scale (DADYS-PI/DADYS-PQ cross-informant scale: $d=0.19$; DADYS-CI/DADYS-CQ cross-informant scale: $d=0.27$).

Regarding the association between the DADYS interviews (total symptom scale) and parent- and child-rated emotion regulation strategies, the results revealed moderate to large positive correlations with maladaptive strategies (DADYS-PI: $0.41 \leq r < .69$; DADYS-CI: $0.56 \leq r < .57$; all $p < .01$) and moderate to large negative correlations with adaptive strategies (DADYS-PI: $-0.67 \leq r < -.32$; DADYS-CI: $-0.50 \leq r < -.44$; all $p < .01$).

Regarding the association between the DADYS interviews (total symptom scale) and externalizing symptoms, we found large to very large correlations with parent- and child-rated ODD symptoms (DADYS-PI: $0.58 \leq r < .84$; DADYS-CI: $0.67 \leq r < .73$; all $p < .01$), moderate to large

Table 3 Discriminant validity of DADYS parent interview (DADYS-PI) and child interview (DADYS-CI) in diagnostic subgroups

diagnosis		DADYS-PI ^a					DADYS-CI ^a				
		n	M	SD	Test statistic	d	n	M	SD	Test statistic	d
ODD/CD	No	336	0.63	0.54	$t(304)=19.03; p < .001$	1.65	335	0.54	0.49	$t(210)=12.22; p < .001$	1.25
	Yes	109	1.45	0.33			109	1.13	0.42		
ADHD	No	378	0.73	0.60	$t(154)=13.15; p < .001$	1.19	377	0.61	0.52	$t(105)=8.21; p < .001$	0.94
	Yes	67	1.40	0.33			67	1.09	0.42		
Any Diagnosis: ODD, CD, ADHD or DD	No	301	0.55	0.52	$t(409)=20.92; p < .001$	1.82	336	0.63	0.54	$t(304)=19.03; p < .001$	1.65
	Yes	144	1.40	0.33			109	1.45	0.33		

Test statistics are based on *t*-tests. *M*=mean. *SD*=standard deviation. *d*=Cohen’s *d*. DADYS=Diagnostic Tool for Affective Dysregulation in Children (Görtz-Dorten & Döpfner, 2021a). ODD=oppositional defiant disorder. CD=conduct disorder. ADHD=attention-deficit/hyperactivity disorder. DD=depressive disorder

^aTotal symptoms scale

correlations with parent- and child-rated CD symptoms (DADYS-PI: $0.41 \leq r < .56$; DADYS-CI: $0.42 \leq r < .52$; all $p < .01$), moderate to large correlations with parent- and child-rated ADHD symptoms (DADYS-PI: $0.47 \leq r < .65$; DADYS-CI: $0.53 \leq r < .62$; all $p < .01$), and large to very large correlations with clinician-rated externalizing symptoms (DADYS-PI: $r = .78$; DADYS-CI: $r = .68$; all $p < .01$). Regarding the association between the DADYS interviews (total symptom scale) and clinician-rated internalizing symptoms, correlations were moderate to large (DADYS-PI: $r = .62$; DADYS-CI: $r = .48$; all $p < .01$).

Finally, regarding parent- and child-rated health-related quality of life, we found moderate to large negative correlations with the total symptom scale of the DADYS interviews (DADYS-PI: $-0.65 \leq r < -.43$; DADYS-CI: $-0.58 \leq r < -.53$; all $p < .01$) and small to moderate negative correlations with the functional impairment scale of the DADYS interviews (DADYS-PI: $-0.35 \leq r < -.20$; DADYS-CI: $-0.46 \leq r < -.43$; all $p < .01$).

Discussion

The aim of the present study was to evaluate the semi-structured clinical DADYS interviews for parents and children in a screened community sample of children with and without AD symptoms as well as in an outpatient sample. The results suggest that both the DADYS-PI and the DADYS-CI are promising and overall reliable and valid interviews for assessing AD in children.

For all items assessing symptoms of AD, we found one factor that provided the best fit to the data, both in the DADYS-PI and the DADYS-CI. Thus, all symptom items were combined into the total (AD) symptom scale. Additionally, the functional impairment scale of the DISYPS-III [21] was added to assess AD-specific functional impairment. As the DADYS-PI further allows for the assessment of a categorical DMDD diagnosis, the DADYS encompasses both the broader conceptualization of AD, implying the proneness to a variety of emotional reactions [2, 3], and the more specific DMDD diagnosis in accordance with the DSM-5 [7]. On a more methodological level, it also allows for both a dimensional assessment (AD symptoms, functional impairment) and a categorical assessment (DMDD) of AD. While a categorical approach might aid the decision-making process regarding the need for treatment, a dimensional approach brings several further advantages, such as a more comprehensive and individual assessment, a more precise assessment of treatment response, and less stigmatization [50, 51]. Finally, the DADYS allows not only for the assessment of AD in both a child and a parent interview, but also in parent, child, and teacher questionnaires.

Thus, the DADYS is able to assess AD both dimensionally and categorically, while other existing instruments are only able to assess AD either dimensionally (ARI,

[18], PROMIS, [19], ERC, [20], CL-ARI, [25], E-SWAN, [27]) or categorically (K-SADS-PL, [26]). Only one other instrument besides the DADYS includes an AD-specific functional impairment scale – the CL-ARI [25]. Moreover, the DADYS is the only comprehensive tool focusing on the broad conceptualization of AD in parent, child, teacher, and clinical ratings, while the other measures offer either parent/child rating (ARI, [18], PROMIS, [19], ERC, [20], CBCL-DP, [22], SDQ-DP, [23, 24]) or clinician rating (CL-ARI, [25], K-SADS-PL, [26], E-SWAN, [27]). Finally, comprehensive analyses of reliability and validity of other instruments assessing AD are still lacking [28]. Taken together, the DADYS offers a variety of advantages which have not yet been covered by any other instrument.

As indicators of reliability, we analyzed the internal consistency and the interrater reliability. Generally, we found acceptable to good internal consistencies for the DADYS interviews based on the samples analyzed, with the only exception being the functional impairment scale of the DADYS-PI. While Haller [25] reported an acceptable internal consistency for their irritability-related functional impairment scale (CL-ARI, $\alpha = 0.75$), interestingly, Thöne, Gortz-Dorten [52] likewise found a lower internal consistency for the functional impairment scale used in the DADYS in a clinical interview assessing ADHD based on a parent interview (ILF-External, DISYPS-III, [21]). As an explanation for this finding, the latter authors argued that the items comprise rather heterogeneous aspects of functional impairment. Particularly in a nonclinical sample such as our screened community sample, if a child is angry, irritable, and moody within his/her own family, this might not be similarly the case in school, with peers, or with other adults. It should be noted that the impairment scale of the CL-ARI assesses functional impairment in three domains (i.e. family, school, and peers) while the impairment scale of the DADYS-PI contains, in addition to these three domains, functional impairment in contact with adults outside the family as well as the patient's strain. Even though one item of the functional impairment scale of the DADYS-PI ("impaired relationships with family members") showed an item-total correlation below $r = .30$, excluding this item did not improve the Cronbach's alpha. Moreover, when interpreting the limited internal consistencies of the functional impairment scales, it should be considered that the sample for these scales consisted only of children with reported AD symptoms, which might have led to an underestimation of reliability and validity. When comparing the good internal consistencies of the DADYS total symptom scales ($\alpha = 0.87-89$) with other clinical measures of DMDD symptoms (CL-ARI, $\alpha = 0.78-0.87$; [25], K-SADS-PL, $\alpha = 0.92$, [53]) and with parent-rated measures of AD (DADYS-PQ,

$\alpha=0.72-0.96$, [14], DADYS-Screen, $\alpha=0.94$, [30]), our data showed mostly comparable internal consistencies.

Regarding interrater reliability, we applied different ICC models for the samples according to the characteristics of the respective data. In both samples, we found moderate to excellent interrater reliability for the total symptom scale and the functional impairment scale in the DADYS interviews. To the best of our knowledge, no study has examined the dimensional interrater reliability of AD symptoms in other clinical interviews. Thus, we were unable to compare the good to excellent interrater reliability of the DADYS total symptom scales (ICC[1,1]=0.94, ICC[2,1]=0.87–0.94) with other AD scales. When comparing the moderate to good interrater reliability of the DADYS functional impairment scales (ICC[1,1]=0.63, ICC[2,1]=0.72–0.85) with clinical measures of functional impairment in externalizing symptoms, we found lower scores than the ILF-External (DISYPS-III, ICC[1,1]=0.89–0.92, [52]), which may be explained by the fact that the ILF-External was evaluated using a clinical sample. Interrater agreement for the dichotomous DMDD diagnoses was substantial to perfect ($\kappa=0.73-1.00$). Compared to other measures of DMDD diagnoses, our values were comparable to slightly higher (K-SADS-PL, $\kappa=0.63$, [54], Conners' Rating Scale, $\kappa=0.68$, [55]).

Taken together, our results on the internal consistency and interrater agreement of the DADYS-PI and the DADYS-CI for the scales (i.e. dimensional assessment) and for the DMDD diagnosis (i.e. categorical assessment) further strengthen our argument that the DADYS can reliably assess AD from both a dimensional and categorical perspective.

As indicators of validity, we analyzed factorial, discriminant, concurrent, convergent, and divergent validity. Regarding factorial validity, almost half of the variance was explained by the total symptom factor. All factor loadings were acceptable (0.38 to 0.88), except for the item exuberance (0.25–0.39). Comparing the factor loadings of the total symptom scales of the DADYS interviews with clinical measures of DMDD (K-SADS-PL, 0.52 to 0.90, [53]) and parent-rated measures of AD (DADYS-Screen, 0.64–0.86, [30], DADYS-PQ, 0.30–1.02, [35]), our data showed mostly comparable factor loadings. In order to capture the broader concept of AD and to be consistent with the DADYS parent questionnaire, we decided to retain the item exuberance.

The present findings further support the concurrent validity of the DADYS interviews: The DADYS interview total symptom scale showed the strongest associations with parent- and child-rated questionnaires of AD with similar item contents, and the DADYS interview functional impairment scale showed the strongest associations with parent- and child-rated measures of

AD-specific functional impairment with the same item content. The associations were particularly strong when using the same informant source – that is the parent interview with the parent questionnaires (AD symptoms: $r=.78-0.87$, functional impairment with DADYS-PQ: $r=.54$) and the child interview with the child questionnaires (AD symptoms: $r=.79$, functional impairment: $r=.67$). The effects are comparable to larger than the associations found between the CL-ARI [25] and clinician-, youth-, and parent-rated measures of irritability ($r=.42-0.89$) and are comparable to the associations between the parent-rated DADYS-Screen [3] and parent- and child-rated measures of AD symptoms ($r=.67-0.83$).

Our analysis of associations with measures of externalizing and internalizing symptoms supports the discriminant, convergent, and divergent validity of the DADYS interviews. First, we found strong discriminative effects of the DADYS interviews for disruptive disorder, ADHD, and any disorder (disruptive, ADHD, and depressive disorder), which were comparable to or larger than the moderate to large discriminative effects found for the DADYS-Screen [3]. Second, we found moderate to very large associations of AD symptoms in the DADYS interview with internalizing and externalizing symptoms. These effects were comparable to or larger than the non-significant to large correlation coefficients reported for internalizing symptoms in the CL-ARI [25] and the moderate to very large correlation coefficients found for internalizing and externalizing symptoms in the DADYS-Screen [4]. When comparing the coefficients within each rater, both DADYS interviews showed the strongest correlations with ODD symptoms, followed by ADHD symptoms, and lastly CD symptoms in all parent-rated measures ($z=2.75-12.31$, $p=.000-0.006$) and in most child-rated measures ($z=1.48-7.46$, $p=.000-0.139$). Furthermore, we found a stronger correlation of externalizing symptoms compared to internalizing symptoms in clinical measures ($z=5.19-5.26$, $p<.001$). These differential relationships are in line with previous research on associations of internalizing and externalizing symptoms both with DMDD symptoms [8, 13] and with AD symptoms (DADYS-Screen, [3]). The associations with both internalizing and externalizing symptoms further emphasize the necessity of the transdiagnostic conceptualization of AD.

Finally, AD symptoms in the DADYS interviews showed small to large associations with emotion regulation strategies and health-related quality of life, which is in line with the associations found with the DADYS-Screen [3].

Besides the evaluation of both the DADYS-PI and the DADYS-CI, it is important to highlight the differences between the two interviews. While AD symptoms showed a very large correlation between the two DADYS

interviews, the correlation for functional impairment was only moderate. Similarly, we found small mean differences for the total scale and moderate differences for the functional impairment scale. As mentioned above, for many associations with validation measures, we found stronger effects when using the same informant source (i.e., parents vs. children). Taken together, our findings suggest that even though there are associations between clinical, parent, and child ratings, there are also some discrepancies among them. Earlier research evaluated these discrepancies as bias or lack of validity, but they are nowadays interpreted as reflecting the unique perspectives of the different informants, with each providing specific and additional information [56]. Thus, an additional strength of both DADYS interviews lies in the possibility to include both parents and children as sources of information in clinical ratings. Furthermore, as a comprehensive assessment tool, the DADYS additionally allows for the inclusion of parent, children and teacher ratings using questionnaires.

Limitations of this study include the restricted age range of 8–12 years within the evaluation, which does not allow for generalizations to children outside of this age range. Furthermore, the screened community sample showed a rather high mean socioeconomic status, which may suggest a lower willingness of parents with low or medium socioeconomic status to participate in the study, potentially reducing the representativeness of the findings. However, the most important limitation of the present study seems to be that the sample did not comprise the full spectrum of AD, but rather consisted of children in the lowest and the highest percentile on the continuum. Interestingly, previous analyses of the psychometric quality of the parent-rated DADYS-Screen [3] have already investigated a potential increase in effects when examining these extreme groups and found only small deviations between correlations for the total sample and the extreme groups. Since a possible overestimation of the reliability and validity of the DADYS interviews cannot currently be ruled out, the presented analyses and results of the DADYS interviews should be interpreted as promising indications of the reliability and validity of the instrument, but replication in further studies with children and adolescents with different levels of AD are needed in order to ultimately establish reliability and validity. Additionally, it should be kept in mind that the internal consistency of the impairment scale of the DADYS-PI must be rated as insufficient and that both the correlation between the symptom scale and the impairment scale of the DADYS-PI and the correlation between the impairment scales of the parent and child interviews were only moderate. In particular, since psychometric studies of related instruments have also found indications of limited psychometric quality of impairment scales

(e.g., [52]), caution is warranted when using the impairment scale of the DADYS-PI, and further research on the functional impairment assessed by parents is needed. For example, it might be worthwhile to investigate whether parental burden has a significant influence on the perceived and reported functional impairment of the child. Lastly, the analyses were based on a cross-sectional design, and it would also be interesting to gain insight into longitudinal associations of child AD with later internalizing and externalizing symptoms (predictive validity, see e.g., [57]). Strengths of this study include the elaborate process of developing the DADYS interviews, the ability of the DADYS to assess AD dimensionally and categorically using the child and the parent as informants, the large screened community sample including children both with and without AD, the additional inclusion of an outpatient sample for interrater reliability, and finally, the inclusion of diverse perspectives for validity analyses (child, parent, and clinical ratings).

Conclusions

This study contributes to the assessment and understanding of children with AD. We evaluated two newly developed semi-structured clinical DADYS interviews—one for parents and one for children. As such, the study is the first to evaluate a clinical interview assessing the broader transdiagnostic conceptualization of AD. Generally, our analyses of internal consistencies and interrater agreement support the reliability of both DADYS interviews. Furthermore, exploratory factor analyses, discriminant analyses, and correlation analyses support the factorial, discriminant, concurrent, convergent, and divergent validity of both DADYS interviews. Since the two DADYS interviews allow for both categorical and dimensional assessment and the inclusion of parents and children as informants, the measure might contribute to the identification of children with AD and the assessment of treatment response.

Abbreviations

AD	Affective dysregulation
ADHD	Attention-deficit/hyperactivity disorder
ARI	Affective Reactivity Index
CBCL-DP	Child Behavior Checklist–Dysregulation Profile
CD	Conduct disorder
CL-ARI	Clinician Affective Reactivity Index
DADYS	Diagnostic Tool for Affective Dysregulation in Children
DADYS-CI	DADYS clinical child interview
DADYS-CQ	DADYS child questionnaire
DADYS-PI	DADYS clinical parent interview
DADYS-PQ	DADYS parent questionnaire
DADYS-Screen	DADYS screening questionnaire
DADYS-TQ	DADYS teacher questionnaire
DISYPS-III	Diagnostic System for Mental Disorders in children and adolescents according to ICD-10 and DSM-5
DMDD	Disruptive mood dysregulation disorder
DSM-5	Diagnostic and Statistical Manual 5th edition
ERC	Emotion Regulation Checklist

E-SWAN	Extended Strengths and Weaknesses Assessment of Normal Behavior
FRUST	Questionnaire for the Regulation of Frustration in Children
ICC	Intraclass correlations
ICC(1,1)	ICC one-way random-effects, absolute agreement model for single rater
ICC(2,1)	ICC two-way random effects, absolute agreement model for single rater
ICD-11	International Classification of Diseases 11th edition
KMO	Kaiser-Meyer-Olkin
K-SADS-PL	Kiddie Schedule for Affective Disorders and Schizophrenia for School Aged Children Present and Lifetime Version
NoAD	Children in the lowest 10% raw scores on the DADYS-Screen
ODD	Oppositional defiant disorder
PCA	Principal component analysis
PFA	Principal factor analysis
PROMIS	Patient-Reported Outcome Measurement Information System
RDoC	Research Domain Criteria
SDQ-DP	Strengths and Difficulties Questionnaire–Dysregulation Profile

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13034-024-00762-8>.

Supplementary Material 1

Acknowledgements

We thank all families participating in our study. Furthermore, we thank following members of the study teams: Maurice Breier, Nina Christmann, Veronika Dobler, Franziska Frenk, Monja Groh, Sarah Hohmann, Nathalie Holz, Anna Kaiser, Josepha Katzmann, Katrin Koppisch, Kristin Kuhnke, Lena Lincke, Anna Michelsen, Sabina Millenet, Christiane Otto, Anne Schreiner, Marie Steiner, Susanne Steinhauser and Matthias Winkler. This trial was supported by the Clinical Trials Centre Cologne (CTCC), Medical Faculty, University of Cologne, by performing the monitoring. For English-language proofreading, we thank Sarah Mannion.

Members of the ADOPT consortium are: Dorothee Bernheim, Stefanie Bienioschek, Maren Boecker, Daniel Brandeis, Kristina Butz, Jörg M. Fegert, Franziska Giller, Carolina Goldbeck, Martin Hellmich, Christine Igel, Michaela Junghänel, Anne Ritschel, Jennifer Schroth, Anne Schüller, Marion Steiner and Anne Uhlmann.

Author contributions

AT conceptualized the data analytic plan, analyzed and interpreted the data, and developed the first draft of the manuscript. Furthermore, she coordinated the ADOPT study. SZL was involved in the development of the DADYS interviews. SZL and CG were involved in the recruitment and data acquisition of the ADOPT study site in Cologne. PA was involved in the recruitment and data acquisition of the ADOPT study site in Mannheim. AD was involved in the recruitment and data acquisition of the ADOPT study site in Rostock. AGD and MD developed the DADYS tool including the DADYS interviews, and designed and coordinated the ADOPT study. MD was additionally involved in the development of the online intervention of the ADOPT Online study. TB, CH, MK, URS, and VR were site leaders for the ADOPT study. CH was the principal investigator of the ADOPT Online study including the development of the online intervention and the coordination of blinded ratings of the DADYS parent interviews. URS and AK were involved in the development of the DADYS tool. All authors critically revised the manuscript for important intellectual content, and all authors gave final approval of the latest version of the manuscript and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding

Open Access funding enabled and organized by Projekt DEAL. The ADOPT Online study was funded by the German Federal Ministry of Education and Research (FKZ 01GL1741D).

Availability of data and materials

Data are available upon reasonable request after the publication of the main results of the ADOPT study.

Declarations

Ethics approval and consent to participate

Approval for the ADOPT Online study was obtained from the ethics committee of the University of Cologne (18–033), the University Hospital of Dresden (EK 35409218), the University Hospital of Mannheim (2018-554 N-MA), and the University Hospital of Ulm (297/18). The procedures used in this study adhere to the tenets of the Declaration of Helsinki. The parents of all participating children and adolescents provided informed consent for participation.

Consent for publication

The parents of all participating children and adolescents provided informed consent for publication.

Competing interests

AGD receives royalties from publishing companies as an author of books and treatment manuals on child behavioral therapy and of assessment manuals, including the treatment manuals evaluated in this trial. She receives income as a consultant for Child Behavior Therapy at the National Association of Statutory Health Insurance Physicians. She also receives consulting income and research support from Medice and eyelevel GmbH. TB served in an advisory or consultancy role for eyelevel GmbH, Infectopharm, Medice, Neurim Pharmaceuticals, Oberberg GmbH, and Takeda. He received conference support or speaker's fees from Janssen, Medice and Takeda. He received royalties from Hogrefe, Kohlhammer, CIP Medien, and Oxford University Press; the present work is unrelated to these relationships. CH receives royalties from a publishing company as the author of a treatment manual. MK receives royalties from publishing companies as an author of books. He served as PI or CI in clinical trials of Lundbeck, Pascoe, and Janssen-Cilag. He served as a scientific advisor for Janssen. The present work is unrelated to the above grants and relationships. VR has received lecture fees from Infectopharm and Medice. He has carried out clinical trials in cooperation with Servier and Shire Pharmaceuticals/Takeda. The present work is unrelated to the above grants and relationships. MD received royalties from publishing companies as an author of books and treatment manuals on child behavioral therapy and of assessment manuals published by Beltz, Elsevier, Enke, Guilford, Hogrefe, Huber, Kohlhammer, Schattauer, Springer, and Wiley. He received income as a consultant for Child Behavior Therapy at the National Association of Statutory Health Insurance Physicians. He also received consulting income and research support from Lilly, Medice, Takeda, and eyelevel GmbH. AT, SZL, CG, LTK, URS, AK, PA, AD, and GK declare no competing interests.

Author details

¹School of Child and Adolescent Cognitive Behavior Therapy (AKiP), Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

²Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

³Department of Child and Adolescent Psychiatry, Psychotherapy, and Psychosomatics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

⁴Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

⁵Faculty of Human Sciences, University of Cologne, Cologne, Germany

⁶Department of Child and Adolescent Psychiatry/Psychotherapy, University Hospital Ulm, Ulm, Germany

⁷Department of Child and Adolescent Psychiatry, Neurology, Psychosomatics, and Psychotherapy, University Medical Center Rostock, Rostock, Germany

⁸Department of Child and Adolescent Psychiatry and Psychotherapy, TUD Dresden University of Technology, Dresden, Germany

Received: 27 November 2023 / Accepted: 4 June 2024

Published online: 20 June 2024

References

- Leibenluft E. Severe mood dysregulation, irritability, and the diagnostic boundaries of bipolar disorder in youths. *Am J Psychiatry*. 2011;168(2):129–42.
- Döpfner M, Katzmann J, Hanisch C, Fegert JM, Kölch M, Ritschel A, et al. Affective dysregulation in childhood - optimizing prevention and treatment: protocol of three randomized controlled trials in the ADOPT study. *BMC Psychiatry*. 2019;19(1):264.
- Treier AK, Döpfner M, Ravens-Sieberer U, Görtz-Dorten A, Boecker M, Goldbeck C, et al. Screening for affective dysregulation in school-aged children: relationship with comprehensive measures of affective dysregulation and related mental disorders. *Eur Child Adolesc Psychiatry*. 2024;33(2):381–90. <https://doi.org/10.1007/s00787-023-02166-z>
- Paulus FW, Ohmann S, Mohler E, Plener P, Popow C. Emotional dysregulation in children and adolescents with psychiatric disorders. *Narrative Rev Front Psychiatry*. 2021;12:628252.
- Legenbauer T, Hubner J, Pinnow M, Ball A, Pniewski B, Holtmann M. Proper emotion recognition, dysfunctional emotion regulation. *Z Kinder Jugendpsychiatr Psychother*. 2018;46(1):7–16.
- Stringaris A, Taylor E. *Disruptive mood: irritability in children and adolescents*. New York: Oxford University Press; 2015.
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 5th ed. Washington, DC: Author; 2013.
- Copeland WE, Angold A, Costello EJ, Egger H. Prevalence, comorbidity, and correlates of DSM-5 proposed disruptive mood dysregulation disorder. *Am J Psychiatry*. 2013;170(2):173–9.
- Mayes SD, Mathiowetz C, Kokotovich C, Waxmonsky J, Baweja R, Calhoun SL, et al. Stability of disruptive mood dysregulation disorder symptoms (irritable-angry mood and temper outbursts) throughout childhood and adolescence in a general population sample. *J Abnorm Child Psychol*. 2015;43(8):1543–9.
- Zero to Three. DC: 0–5: Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood. revised ed. Washington, DC: Zero to Three Press; 2016.
- Bödeker K, Watrin-Avino LM, Martin A, Schlenso-Schuster F, Janssen M, Friese L, et al. Assessment and Diagnostic classification using DC:0–5 in early childhood Mental Health clinics: the protocol for the Developmental Psychiatry Diagnostic challenges Study (DePsy). *Children (Basel)*. 2023;10(11):1770.
- World Health Organization. *International Statistical Classification of Diseases and related health problems, 11th revision (ICD-11)*. Geneva: World Health Organization; 2019.
- Bruno A, Celebre L, Torre G, Pandolfo G, Mento C, Cedro C, et al. Focus on disruptive mood dysregulation disorder: a review of the literature. *Psychiatry Res*. 2019;279:323–30.
- Junghänel M, Thöne A-K, Ginsberg C, Görtz-Dorten A, Frenk F, Mücke K, et al. Irritability and emotional impulsivity as core feature of ADHD and ODD in children. *J Psychopathol Behav Assess*. 2022;44(3):679–97.
- Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S, et al. The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clin Psychol Sci*. 2014;2(2):119–37.
- Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research Domain Criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry*. 2010;67(7):748–51.
- Meyers E, DeSerisy M, Roy AK. Disruptive mood dysregulation disorder (DMDD): an RDoC perspective. *J Affect Disord*. 2017;216:117–22.
- Stringaris A, Goodman R, Ferdinando S, Razdan V, Muhrer E, Leibenluft E, et al. The affective reactivity index: a concise irritability scale for clinical and research settings. *J Child Psychol Psychiatry*. 2012;53(11):1109–17.
- Irwin DE, Stucky BD, Langer MM, Thissen D, DeWitt EM, Lai J-S, et al. PROMIS Pediatric anger scale: an item response theory analysis. *Qual Life Res*. 2012;21(4):697–706.
- Shields A, Cicchetti D. Emotion regulation among school-age children: the development and validation of a new criterion Q-sort scale. *Dev Psychol*. 1997;33(6):906–16.
- Döpfner M, Görtz-Dorten A. Diagnostik-System für psychische Störungen Nach ICD-10 und DSM-5 für Kinder Und Jugendliche - III. Bern: Hogrefe; 2017.
- Althoff RR. Dysregulated children reconsidered. *J Am Acad Child Adolesc Psychiatry*. 2010;49(4):302–5.
- Holtmann M, Becker A, Banaschewski T, Rothenberger A, Roessner V. Psychometric validity of the strengths and difficulties questionnaire-dysregulation profile. *Psychopathology*. 2011;44(1):53–9.
- Deutz MHF, Shi Q, Vossen HGM, Huijding J, Prinzie P, Deković M, et al. Evaluation of the strengths and difficulties Questionnaire-Dysregulation Profile (SDQ-DP). *Psychol Assess*. 2018;30(9):1174–85.
- Haller SP, Kircanski K, Stringaris A, Clayton M, Bui H, Agorsor C, et al. The Clinician Affective Reactivity Index: validity and reliability of a clinician-rated assessment of irritability. *Behav Ther*. 2020;51(2):283–93.
- Kaufman J, Birmaher B, Brent D, Rao UMA, Flynn C, Moreci P, et al. 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*. 1997;36(7):980–8.
- Alexander L, Salum G, Swanson J, Milham M. Measuring strengths and weaknesses in dimensional psychiatry. *J Child Psychol Psychiatr*. 2020;61:40–50.
- Murner-Lavanchy I, Kaess M, Koenig J. Diagnostic instruments for the assessment of disruptive mood dysregulation disorder: a systematic review of the literature. *Eur Child Adolesc Psychiatry*. 2023;32(1):17–39.
- Görtz-Dorten A, Döpfner M. Diagnostikum für Affektive Dysregulationen bei Kindern. [Unpublished manuscript]. In press. 2021.
- Otto C, Kaman A, Barkmann C, Döpfner M, Görtz-Dorten A, Ginsberg C et al. The DADYS-Screen. Development and evaluation of a screening tool for affective dysregulation in children. *Assessment*. 2022.
- Lidzba K, Christiansen H, Drechsler R, Conners 3 - Conners Skalen zu Aufmerksamkeit Und Verhalten 3 (Manual). 3 ed. Bern: Huber; 2013.
- Döpfner M, Plück J, Kinnen C, Arbeitsgruppe Deutsche Child Behavior Checklist, von Thomas M. Achenbach. Göttingen: Hogrefe; 2014.
- McQuillan ME, Kultur EC, Bates JE, O'Reilly LM, Dodge KA, Lansford JE, et al. Dysregulation in children: origins and implications from age 5 to age 28. *Dev Psychopathol*. 2018;30(2):695–713.
- Görtz-Dorten A, Döpfner M. Fragebogen zur Regulation unangenehmer Stimmungen von Kindern. [Unpublished manuscript]. In press. 2021.
- Junghänel M, Wand H, Dose C, Thöne AK, Treier AK, Hanisch C, et al. Validation of a new emotion regulation self-report questionnaire for children. *BMC Psychiatry*. 2022;22(1):820.
- IBM Corp. IBM SPSS statistics for Windows (Version 29.0) [Computer software]. Armonk: IBM Corp; 2022.
- Bennett D. How can I deal with missing data in my study? *Aust N Z J Public Health*. 2001;25:464–9.
- Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale: Lawrence Erlbaum Associates; 1988.
- Cattell RB. The scree test for the number of factors. *Multivar Behav Res*. 1966;1(2):245–76.
- Velicer WF. Determining the number of components from the matrix of partial correlations. *Psychometrika*. 1976;41(3):321–7.
- Horn JL. A rationale and test for the number of factors in factor analysis. *Psychometrika*. 1965;30:179–85.
- Hair JF, Black WC, Babin BJ, Anderson RE. *Multivariate data analysis*. 7th ed. Upper Saddle River: Pearson Education; 2014.
- Nunnally JC. *Psychometric theory*. 2nd ed. New York: McGraw-Hill; 1978.
- Field A. *Discovering statistics using IBM SPSS statistics*. 5th ed. London: Sage Publications; 2018.
- Shrout P, Fleiss J. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420–8.
- Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–63.
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37–46.
- Landis J, Koch G. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–74.
- Lampert T, Hoebel J, Kuntz B, Müters S, Kroll LE. Messung Des sozioökonomischen Status Und Des Subjektiven Sozialen Status in KiGGS Welle 2. *J Health Monit*. 2018;3.
- Lahey BB, Tiemeier H, Krueger RF. Seven reasons why binary diagnostic categories should be replaced with empirically sounder and less stigmatizing dimensions. *JCPP Adv*. 2022;2(4):e12108.
- Kraemer H. DSM categories and dimensions in clinical and research contexts. *Int J Methods Psychiatr Res*. 2007;16:58–15.
- Thöne AK, Görtz-Dorten A, Altenberger P, Dose C, Geldermann N, Hautmann C, et al. Toward a dimensional assessment of externalizing disorders in

- children: reliability and validity of a semi-structured parent interview. *Front Psychol.* 2020;11:1840.
53. de la Pena FR, Rosetti MF, Rodriguez-Delgado A, Villavicencio LR, Palacio JD, Montiel C, et al. Construct validity and parent-child agreement of the six new or modified disorders included in the Spanish version of the Kiddie schedule for affective disorders and Schizophrenia present and Lifetime Version DSM-5 (K-SADS-PL-5). *J Psychiatr Res.* 2018;101:28–33.
54. Unal F, Oktem F, Cetin Cuhadaroglu F, Cengel Kultur SE, Akdemir D, Foto Ozdemir D, et al. Reliability and validity of the schedule for affective disorders and Schizophrenia for School-Age Children-Present and Lifetime Version, DSM-5 November 2016-Turkish adaptation (K-SADS-PL-DSM-5-T). *Turkish J Psychiatry.* 2019;30:42–50.
55. Tufan E, Topal Z, Demir N, Taskiran S, Savci U, Cansiz MA, et al. Sociodemographic and clinical features of disruptive mood dysregulation disorder: a chart review. *J Child Adolesc Psychopharmacol.* 2016;26(2):94–100.
56. De Los Reyes A, Augenstein TM, Wang M, Thomas SA, Drabick DAG, Burgers DE, et al. The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychol Bull.* 2015;141(4):858–900.
57. Kessel EM, Dougherty LR, Kujawa A, Hajcak G, Carlson GA, Klein DN. Longitudinal associations between preschool disruptive mood dysregulation disorder symptoms and neural reactivity to monetary reward during preadolescence. *J Child Adolesc Psychopharmacol.* 2016;26(2):131–7.
58. Ravens-Sieberer U, Erhart M, Gosch A, Wille N, European KG. Mental health of children and adolescents in 12 European countries - results from the European KIDSCREEN study. *Clin Psychol Psychother.* 2018;15(3):154–63. <https://doi.org/10.1002/cpp.574>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.