

# Supplementary scales for the school-age forms of the Achenbach System of Empirically Based Assessment rated by adolescents, parents, and teachers: Psychometric properties in German samples

Julia Plück<sup>1,2\*</sup>, Laurence Nawab<sup>1,2</sup>, Elena Kamenetzka<sup>1,2</sup>, Manfred Döpfner<sup>1,2</sup>

<sup>1</sup>Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Cologne

<sup>2</sup>School for Child and Adolescent Cognitive Behavior Therapy (AKiP), University Hospital Cologne

\*Corresponding author: julia.plueck@uk-koeln.de

## Abstract

**Background:** Based on Achenbach's school-age questionnaires, research groups have investigated supplementary scales for stress problems, obsessive-compulsive problems, sluggish cognitive tempo, positive qualities, dysregulation, autism spectrum disorders, and mania in 6-18-year-olds partly only in some of the three perspectives the Achenbach System of Empirically Based Assessment (ASEBA) provides.

**Objective:** We aimed to evaluate these dimensions for the German-language forms and, if possible, to extend their use to further rating perspectives.

**Methods:** The internal consistencies of the supplementary scales were examined for three types of informants (parents, adolescents, and teachers) and different samples (community sample, clinical sample, and disorder-specific subsamples). Age- and gender-specific effects are displayed as well as cross-informant correlations. Additionally, different aspects of validity were analyzed: (a) convergent/divergent validity via correlations with traditional ASEBA scales (problem scales as well as the scales oriented to the Diagnostic and Statistical Manual of Mental Disorders 5th edition); (b) discriminative validity via differences between clinical vs. community-based sample as well as disorder-specific subgroup vs. clinical sample.

**Results:** Most of the supplementary scales showed at least acceptable internal consistency. For some scales, we found significant but rather small and informant-dependent gender and age differences. Convergent validity of the supplementary scales differed across informants. Mean differences between the supplementary scales in the clinical and the community sample as well as the diagnosis-specific subsamples were mostly significant, with predominantly large effect sizes.

**Conclusions:** Overall, the validity and reliability of the supplementary scales differed depending on informants and subgroups. While further research is necessary before the supplementary scales are implemented in clinical practice, initial recommendations for their use are derived.

**Trial registration:** This project was carried out as a reanalysis of the datasets upon which the German norms for the school-age versions are based (1). Therefore, the trial was not registered.

**Keywords:** ASEBA, German school-age versions, 2007 scales, supplementary scales, psychometric properties.

## Introduction

The Achenbach System of Empirically Based Assessment (ASEBA) represents the most widely used and cross-culturally established rating scales internationally (3-4). The questionnaires for school-age children and adolescents are also available in a German version for all three informant perspectives: parent report (CBCL/6-18R), youth self-report (YSR/11-18R), and teacher report (TRF/6-18R) (1). These are well established in Germany in child- and adolescent psychiatry as well as psychotherapy for inpatient and outpatient facilities and practices,

recommended as basic instruments for broadband multimodal diagnostic of behavioral and emotional problems (2). Factor analyses generated eight empirically based cross-informant problem scales (Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior, and Aggressive Behavior) (4-7). Additionally, six DSM-oriented scales were developed, which are based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-5: Depressive Problems, Anxiety Problems, Somatic

Problems, Attention Deficit/Hyperactivity Problems, Oppositional Defiant Problems, and Conduct Problems (8)). The scale Attention Deficit/Hyperactivity Problems in the teacher report version is further divided into two subscales: Inattention and Hyperactivity-Impulsivity (3). Expanding on these “2001 scales”, Achenbach and Rescorla integrated the following four supplementary “2007 scales” into their Multicultural Supplement to the Manual for the ASEBA School-Age Forms & Profiles (9): Based on previous findings, Achenbach & Rescorla proposed a 14-item scale as a supplement to the CBCL, initially named Posttraumatic Stress Problems scale (PTSP) (9-11) and renamed as the Stress Problems Scale (SPS) by Achenbach, Rescorla & Ivanova in 2015 (9, 10) to reflect that the scale measures general stress rather than specific posttraumatic stress symptoms (12, 13). Furthermore, a CBCL Obsessive-Compulsive Problems scale (OCP) (8, 13) and a TRF Sluggish Cognitive Tempo scale (SCT)(9, 15) were developed and implemented in the Multicultural Supplement to the Manual for the ASEBA School-Age Forms & Profiles (9). Besides these problem-oriented scales, a fourth “2007 scale” was generated: the YSR Positive Qualities Scale (PQS) (9). The PQS covers 14 positive items from the YSR (9), with significantly higher scores being found in girls compared to boys and in older compared to younger adolescents (16). More recently, other research groups have developed the following additional supplementary scales, which have not yet been integrated into the ASEBA family: CBCL and TRF Dysregulation Scale (DRS1) (17,18) as well as a weighted and rounded Dysregulation Scale (DRS2) (19), different types of CBCL and TRF Autism Spectrum Disorder scales (20, 22) and the CBCL Mania Scale (23).

In 1995, a dysregulation profile was identified, which is characterized by high scores on the following rating scales of the 1991 CBCL version: Aggressive Behavior, Anxious/Depressed, and Attention Problems (18). Other researchers (17) calculated a sum score from the 41 items of these three subscales (2001 version of the CBCL). Rescorla et al. (24) conducted a cross-cultural comparison of the sum score of the dysregulation profile and found mostly minor differences between gender and age across informants. Based on the previous research, in 2018, McQuillan et al. (19) suggested a weighted and rounded score in order to improve the conceptual precision.

Furthermore, two data-driven specific autism spectrum disorder (ASD) scales were constructed - a 9-item CBCL Autism spectrum disorder scale (ASD1) (20) and a 10-item TRF Autism spectrum disorder scale (ASD2) (21) each developed both in community samples and in heterogeneous clinical

samples. The two scales differ in three items. More recently, a third, data-driven ASD scale with 15 items was developed, which overlaps in six items with the ASD1 and in nine items with the ASD2 (22).

Overall, in most analyses, these supplementary scales have shown at least acceptable internal consistencies ( $\geq .70$ ) for the informant perspectives analyzed in the respective studies (9, 13, 14, 19, 21-23, 25-30). Moreover, validity has been demonstrated for several of the scales. While low correlations between the respective informants were reported for SCT (CBCL and YSR) and OCP, correlations were moderate for SPS ( $r = .38$ ) and even stronger for DRS ( $r > .50$ ) (9, 24). Further validity analyses revealed that very high SPS scores in the YSR were associated with a moderate increase in the likelihood of a posttraumatic stress disorder diagnosis (25). Additionally, children and adolescents with a diagnosis of obsessive-compulsive disorder obtained significantly higher scores on the OCP scale compared to a psychiatric control sample and a population control sample (14).

To the best of our knowledge, no study has been conducted examining the supplementary scales of the Achenbach system for school-age across informants in a community sample, a clinical sample, and disorder-specific subsamples for aspects of their convergent/divergent and discriminant validity as well as their internal consistencies. This is the aim of the current analyses. In view of the already confirmed intercultural reliability of the instruments, we hope our analyses may contribute to the canon of results on these instruments. Interculturally validated instruments enable the transferability of results of both epidemiological and clinical-interventional studies to the target group. In addition, such instruments can also be used reliably for patients with a migration background, who in modern societies have their natural share among those seeking advice from help facilities.

## Methods

### Sample

The community sample as well as the clinical sample in use correspond to the samples that were also used for manual of the German version of CBCL/YSR and TRF in 2014 (1). The data of the community sample encompassed 2,471 parent reports ( $M_{\text{age}} = 12.3$ ;  $SD_{\text{age}} = 3.3$ ; 50.5% male) and 1,797 self-reports ( $n_{\text{PQS}} = 1,791$ ;  $M_{\text{age}} = 13.9$ ;  $SD_{\text{age}} = 2.2$ ; 51.4% male). Both perspectives, CBCL and YSR, were available for 1,757 cases. Teachers' reports had not been collected in the general population sample as this was a household-based survey. Six participants missed more than 10% of the items on the Positive Qualities Scale (YSR), and were not included in the scale-specific analyses. The clinical sample originated

from pre-treatment assessments at the outpatient psychotherapy unit of the School of Child and Adolescent Cognitive Behaviour Therapy at the University of Cologne (AKiP). It comprised children and adolescents aged 6-18 years who were referred to the outpatient clinic for child and adolescent psychotherapy for a variety of psychopathologies. A more detailed description of the distribution of diagnoses within the clinical sample can be found in the (1). Parent reports (N = 1,288; Mage = 11.5; SDage = 3.3; 64.7% male), self-reports (N = 755; Mage = 14.0; SDage = 2.1; 56.4% male), and teacher reports (N = 836; Mage = 10.9; SDage = 3.1; 69.9% male) were included. Both the CBCL and the YSR were available in 705 cases, the CBCL and TRF in 799 cases, and the TRF and YSR in 396 cases. In all data collections, caregivers and patients agreed to data processing for group statistical analyses.

### Measures

The items of CBCL, TRF and YSR are rated on a three-point scale (0 = not true [as far as you know], 1 = sometimes or somewhat true, or 2 = very true or often true) (1). The CBCL and TRF was completed by the caregivers and teachers of children and adolescents aged 6-18 years and the YSR was completed by adolescents aged 11-18 years. As defined by the standards of the ASEBA questionnaires, parents and adolescents assessed the behavioral problems of the last six months while teachers considered the last two months (3). Following Achenbach and Rescorla (3), a maximum of eight items missing from the total score was permitted, with missing values replaced with "0". All possibly available items of each supplementary scale were considered for the analyses. The three items that were necessary for at least one supplementary scale (4: fails to finish, 5: Enjoys little and 78: inattentive or easily distracted) due to the version adaptation of the questionnaires were not available. These items were replaced by the mean score achieved for the items of the respective problem scale for each case in both datasets. The following supplementary scales were analyzed across the respective informants and samples: Stress Problems Scale (SPS) (9), Obsessive-Compulsive Problems scale (OCP) (9), Sluggish Cognitive Tempo scale (SCT) (available for the CBCL and TRF) (9), and the YSR Positive Qualities scale (PQS) (9). Furthermore, we investigated the summed Dysregulation Scale (DRS1) (17), the weighted and rounded Dysregulation Scale (DRS2) (19), three Autism Spectrum Disorder scales according to Ooi et al. (ASD1) (20), So et al. (ASD2) (21), and Offermans et al. (ASD3) (22), as well as the Mania Scale (MAS) (23).

### Data Analyses

Analyses were performed using IBM SPSS Statistics, Versions 27 and 28 (www.spss.com). First, internal consistencies (Mc Donald's  $\Omega$ /Cronbach's  $\alpha$  for Comparisons with other studies) were calculated with values  $\geq .90$  considered as excellent, values from .80 to .89 as good, values from .70 to  $\approx .79$  as adequate, and values  $< .70$  as inadequate (30). For comparison purposes with existing studies, however, the Cronbach's alpha values are also presented in the supplement (table S1). Furthermore, item characteristics (means and standard deviations of items, item-total correlations) were computed for each supplementary scale (Tables S1 to S14 in the supplement). As recommended by Achenbach and Rescorla (3), for all further analyses, raw scores of the scales were used. As some age and gender effects as well as low cross-informant correlations of corresponding scales had been found in the previous analyses of the traditional problem scales, Pearson (and additionally Spearman, see supplement) correlations were calculated to analyze effects of age (in years) and gender (male=1, female=2). According to Cohen (1988)- (32), correlations from  $r=.10$  to  $r=.29$  are considered as low, from  $r=.30$  to  $r=.49$  as moderate, and  $\geq r=.50$  as high. For main areas confidence intervals (95%) of the correlations are reported to prevent overinterpretation given the rather large sample size. Second, convergent/divergent validity was analyzed by Pearson (and additionally Spearman, see supplement) correlations between the supplementary scales and the traditional empirical scales as well as the DSM-oriented scales (3). To enable comparisons of correlations with Internalizing Problems (INT) and Externalizing Problems (EXT), we used an online calculation tool (32) providing a formula for dependent groups with a third variable based on z-transformations (34-36). Third, to provide insights into the discriminant validity of the scales, mean differences between the clinical and community samples for the CBCL and YSR were calculated. As homogeneity of variance was largely lacking, the Welch test was chosen. Effect sizes of means (Cohen's d:  $d=0.20$  small effect size;  $d=0.50$  medium effect size,  $d=0.80$  large effect size) (32), were calculated. Additionally, for two areas (obsessive-compulsive disorder for OCP, pervasive developmental disorders for ASD1, ASD2, and ASD3) mean differences between the clinical subgroup characterized by the nearest respective diagnosis (as far as possible) and the rest of the clinical sample were calculated.

**TABLE 1.** Scale scores and internal consistencies in clinic and community sample across informants

	CBCL/6-18			YSR/11-18			TRF/6-18		
	Origin N	Clinic 1288	Community 2471	Clinic 755	Community 1797	Clinic 836			
Age (years)									
M		11.5	12.3	14.0	13.9	10.9			
(SD)		(3.3)	(3.3)	(2.1)	(2.2)	(3.1)			
Male%		64.7	50.5	56.4	51.4	69.9			
Scale	(n <sub>it</sub> ) Ω	r <sub>it</sub> (Range)	M SD	(n <sub>it</sub> ) Ω	r <sub>it</sub> (Range)	M SD	(n <sub>it</sub> ) Ω	r <sub>it</sub> (Range)	M SD
SPS	(14) .72	(.18-.50)	9.07 4.56	(14) .79	(.23-.61)	8.89 5.02	(13) .69	(-.21-.53)	4.57 3.81
OCP	(8) .69	(.26-.51)	3.18 2.86	(8) .74	(.33-.55)	3.92 3.37	(8) .66	(.26-.45)	1.56 1.96
SCT <sup>1,2</sup>	(4) .66	(.34-.54)	1.79 1.77	(4) .50	(.27-.33)	0.51 0.91	(5) .81	(.47-.66)	17.09 6.19
PQS <sup>3</sup>				(14) .73	(.14-.47)	18.83 4.37	(62) .95 <sup>A</sup>	(-.02-.71)	12.71 9.51
DRS1	(41) .89	(.14-.60)	24.32 12.22	(39) .89	(.20-.63)	21.33 11.34	(9) .63	(.15-.48)	29.08 19.52
DRS2			24.30 11.98			22.11 11.61	(10) .74	(.29-.53)	2.77 2.67
ASD1	(9) .58	(.13-.43)	3.56 2.81	(9) .51	(.12-.31)	3.57 2.70	(10) .74	(.29-.53)	1.65 1.89
ASD2	(10) .64	(.20-.43)	3.99 3.13	(9) .58	(.22-.39)	3.97 2.88	(15) .80	(.22-.59)	1.82 1.99
ASD3 <sup>1</sup>	(15) .78	(.21-.55)	6.76 4.87	(14) .72	(.18-.45)	6.69 4.37	(15) .87	(.03-.73)	3.40 3.30
MAS <sup>1</sup>	(19) .79	(.13-.59)	8.26 5.25	(17) .75	(.19-.46)	7.94 4.78	(15) .87	(.03-.73)	4.93 4.13

Notes. CBCL/6-18 = Parent reports (6;0 to 18;11 years); YSR/11-18 = Self reports (11;0 to 18;11 years); TRF/6-18 = Teacher reports (6;0 to 18;11 years); n<sub>it</sub> = Number of items; Ω = McDonald's Omega; r<sub>it</sub> (Range) = range of internal consistency; M = Mean scale score; SD = Standard Deviation; SPS = Stress Problems scale (Achenbach & Rescorla, 2007); OCP = Obsessive-Compulsive Problems scale (Achenbach & Rescorla, 2007); SCT = Sluggish Cognitive Tempo scale (Achenbach & Rescorla, 2007); Inew generated items for the analysis of internal consistencies excluded; 2scale not available in YSR/11-18; PQS = Positive Qualities scale (Achenbach & Rescorla, 2007); 3scale not available in CBCL/6-18 and TRF/6-18 & reduced community sample size for YSR (nPQS = 1791); DRS1 = Dysregulation scale (Ayer et al., 2009); DRS2 = Weighted and rounded Dysregulation scale (McQuillan et al., 2018); ASD1 = Autism Spectrum Disorder Scale (Ooi et al., 2011); ASD2 = Autism Spectrum Disorder Scale (So et al., 2012); ASD3 = Autism Spectrum Disorder Scale (Offermans et al., 2022); MAS = Mania Scale (Papachristou et al., 2013). A Omega could not be estimated due to item covariances that are negative or zero. Cronbach's Alpha is reported as a substitute.

## Results

### *Internal consistency*

As shown in table 1, for four of the eight supplementary scales analyzed in the CBCL clinical sample, internal consistency was at least adequate (SPS, ASD3, MAS:  $.73 \leq \Omega \leq .79$ ) or good (DRS1, DRS2:  $\Omega = .89$ ). By contrast, the other four scales did not show adequate internal consistency across all samples (OCP, SCT, ASD1, ASD2:  $.58 \leq \Omega \leq .69$ ). In the CBCL community sample, the supplementary scales showed very similar internal consistencies to those in the clinical sample ( $.50 \leq \Omega \leq .89$ ). For the YSR, internal consistencies were at least acceptable for all scales and in both samples, with the exception of ASD1 and ASD2 in the clinical and community samples and OCP in the community sample. The PQS showed adequate to good internal consistency in both samples. In the clinical TRF sample, the internal consistencies for SPS, ASD1 and OCP were inadequate ( $\Omega < .70$ ), while ASD2 showed adequate consistency. Furthermore, SCT, ASD3, and MAS showed good internal consistencies ( $.80 \leq \Omega \leq .87$ ), and internal consistency was excellent for both DRS1 and DRS2.

### *Age- and gender-specific correlations*

Table 2 presents the correlations of each supplementary scale with age and gender within both available settings (clinical, community). Nearly all correlations were significant due to the sample size; however, the majority of correlations were low, although some were substantial ( $|\text{CIr}| \leq .30$ ). In the clinical YSR sample, SPS and OCP correlated with age and gender. Additionally, ASD2 and ASD3 correlated positively with age. Furthermore, both DRS scales correlated negatively with gender in the clinical TRF sample, while MAS correlated negatively with gender in the clinical CBCL as well as TRF samples. A look at table S16 shows, that assessment based on the CIs of Spearman's rho did not change these findings.

### *Cross-informant correlations*

Table 3 presents cross-informant correlations in the clinical and community samples. At least moderate correlations emerged between the CBCL and YSR in the community sample ( $.30 \leq \text{CIr} \leq .50$ ), low ones ( $.10 \leq \text{CIr} \leq .30$ ) for most scales in the clinical sample. Between the TRF and CBCL in the clinical sample moderate correlations were found for MAS and ASD3, all other correlations were low. The TRF and YSR showed low but significant correlations for OCP, ASD1 and MAS including the CIs. A look at table S17 shows, that assessment based on the CIs of Spearman's rho did not weaken these findings, the correlations here trended even higher.

### *Correlations between original school-age scales and supplementary scales*

Tables S18.1, S19.1 and S20.1 of the supplement show the Pearson correlations of Achenbach's original school-age scales with the supplementary scales for each rating perspective. Most correlations were statistically significant ( $p \leq .01$ ). With the exception of PQS (YSR only) and SCT for nearly (CBCL and TRF only), all supplementary scales correlated strongly with at least the INT or EXT scale for each perspective and each sample ( $\text{CIr} / \text{CIroh} \geq .50$ ). It should be noted, however, that there is a higher or lower degree of overlap between the items in the respective pairs of scales, which necessarily increases correlations. Furthermore, in the CBCL and TRF clinical samples, the OCP was the only supplementary scale that did not correlate strongly with the Total Problems Scale (TOT). The correlations for each supplementary scale with INT and EXT, respectively, after checking the statistical significance of their differences (see table S21.1 and S21.2), revealed the following picture with only single exceptions when rated based on Spearman's rho instead (see table S18.2, S19.2 and S20.2) and taking CIr/ CIroh into account:

Across informants and samples, SPS showed at least moderate correlations with INT and EXT as well as with the DSM-oriented scales. The correlations with INT were higher than those with EXT. Based on Spearman's rho within the community sample for CBCL correlations with INT and EXT did not differ significantly. The supplementary scale OCP showed stronger significant correlations with INT than with EXT for all informants and samples. Across the samples, OCP showed strong correlations with Depressive Problems and Anxiety Problems. However, the results were less congruent concerning correlations with the further DSM-oriented scales, with low to high effects emerging depending on the respective sample and informant. SCT was strongly correlated with INT across informants. The correlations between SCT and the DSM-oriented scales in both CBCL samples were at least low ( $r = .10$ ). SCT showed at least moderate correlations with the Inattention subscale and a strong correlation with Depressive Problems. Overall, in line with expectation, PQS showed the lowest correlation with the problem-focused scales, with small effect sizes in both YSR samples. Therefore, no significant difference between the correlations with INT or EXT was found. Both DRS correlated strongly with EXT and INT across almost all informants and samples ( $r \geq .50$ ). However, for the CBCL and the TRF, correlations were higher with EXT than with INT, and both scales showed almost the same results

**TABLE 2.** Pearson Correlations of supplement scales with age and gender within both samples for each informant

Scale	CBCL/6-18R			YSR/11-18R			TRF/6-18R		
	Origin	Clinic	Community	Clinic	Community	Clinic	Community	Clinic	Community
	N	1288	2471	755	1797	836			
	n <sub>female</sub>	455	1224	326	874	251			
	n <sub>male</sub>	833	1247	429	923	585			
		r	r	r	r	r	r	r	r
		Cl <sub>r</sub>	Cl <sub>r</sub>	Cl <sub>r</sub>	Cl <sub>r</sub>	Cl <sub>r</sub>	Cl <sub>r</sub>	Cl <sub>r</sub>	Cl <sub>r</sub>
SPS	Age	.06	[-.00, .11]	.21*	[-.14, .28]	.06	[-.01, .10]	.07	[-.01, .14]
	Gender <sup>1</sup>	.06	[<.01, .12]	.29*	[.22, .35]	.08*	[.04, .13]	-.08	[-.15, -.02]
OCP	Age	.11*	[.05, .16]	.27*	[.20, .33]	.11*	[.06, .15]	.06	[<-.01, .13]
	Gender	.12*	[.07, .18]	.32*	[.26, .38]	.07*	[.02, .11]	.02	[-.05, .08]
SCT <sup>2</sup>	Age	.09*	[<.01, .18]	<.01	[<.06, .05]	-		.08	[.01, .14]
	Gender	<.01	[<.05, .06]	-.02	[<.06, .02]	-		-.09*	[<.16, -.02]
PQS <sup>3</sup>	Age	-		-		.12*	[.08, .17]	-	
	Gender	-		-.04	[<.11, .03]	.07*	[.03, .12]	-	
DRS1	Age	-.10*	[<.15, -.05]	-.04	[<.08, <.01]	.06	[.01, .10]	-.11*	[<.17, -.04]
	Gender	-.15*	[<.20, -.09]	-.07*	[<.11, -.03]	.04	[<.01, .08]	-.27*	[<.33, -.20]
DRS2	Age	-.09*	[<.14, -.03]	-.04	[<.08, <.01]	.04	[<.01, .09]	-.09*	[<.16, -.02]
	Gender	-.15*	[<.20, -.10]	-.07*	[<.11, -.03]	.04	[<.01, .08]	-.24*	[<.31, -.18]
ASD1	Age	.04	[<.01, .10]	-.02	[<.06, .02]	.18*	[.11, .25]	-.01	[<.08, .06]
	Gender	-.06	[<.12, -.01]	-.01	[<.05, .03]	.15*	[.08, .22]	-.14*	[<.21, -.07]
ASD2	Age	.07*	[.02, .13]	.01	[<.03, .05]	.24*	[.18, .31]	.01	[<.05, .08]
	Gender	-.03	[<.08, .03]	-.03	[<.07, .01]	.18*	[.11, .25]	-.11*	[<.18, -.04]
ASD3	Age	.08*	[.03, .13]	-.01	[<.05, .03]	.27*	[.20, .34]	.04	[<.03, .11]
	Gender	.01	[<.05, .07]	<.01	[<.04, .04]	.20*	[.13, .27]	-.08	[<.15, -.01]
MAS	Age	-.20*	[<.25, -.14]	-.07*	[<.11, .03]	.07	[<.01, .14]	-.15*	[<.21, -.08]
	Gender	-.25*	[<.30, -.19]	-.10*	[<.14, -.06]	.07	[<.01, .14]	-.28*	[<.34, -.22]

Notes. CBCL/6-18 = Parent reports (6;0 to 18;11 years); YSR/11-18 = Self reports (11;0 to 18;11 years); TRF/6-18 = Teacher reports (6;0 to 18;11 years); n<sub>male</sub> = male sample; Cl<sub>r</sub> = 95% Confidence interval; <sup>1</sup>Gender: 1= male, 2= female; SPS = Stress Problems scale (Achenbach & Rescorla, 2007); OCP = Obsessive-Compulsive Problems scale (Achenbach & Rescorla, 2007); SCT = Sluggish Cognitive Tempo scale (Achenbach & Rescorla, 2007); <sup>2</sup>scale not available in YSR/11-18; <sup>3</sup>scale not available in CBCL/6-18 and TRF/6-18 & reduced community sample size for YSR (n<sub>mas</sub> = 1791); DRS1 = Dysregulation scale (Ayer et al., 2009); DRS2 = Weighted and rounded Dysregulation scale (McQuillan et al., 2018); ASD1 = Autism Spectrum Disorder Scale (Ooi et al., 2011); ASD2 = Autism Spectrum Disorder Scale (So et al., 2012); ASD3 = Autism Spectrum Disorder Scale (Offermans et al., 2022); MAS = Mania Scale (Papachristou et al., 2013); \* p ≤ .01.

**TABLE 3.** Cross-informant Pearson correlations in community and clinic sample.

N	Community				Clinic			
	CBCLxYSR		CBCLxYSR		CBCLxTRF		TRFxYSR	
	1757		705		799		396	
Scale	<i>r</i>	<i>Cl<sub>r</sub></i>	<i>r</i>	<i>Cl<sub>r</sub></i>	<i>r</i>	<i>Cl<sub>r</sub></i>	<i>r</i>	<i>Cl<sub>r</sub></i>
SPS	.50	[.47,.54]	.35	[.28,.41]	.20	[.13,.27]	.09 <sup>+</sup>	[-.01,.18]
OCP	.35	[.31,.39]	.41	[.35,.47]	.20	[.14,.27]	.21	[.12,.30]
SCT <sup>1</sup>	-		.33	[.26,.39]	-		-	
DRS1	.53	[.50,.57]	.34	[.27,.40]	.35	[.29,.41]	.15	[.05,.24]
DRS2	.53	[.48,.56]	.33	[.26,.39]	.33	[.27,.39]	.14	[.04,.23]
ASD1	.39	[.35,.43]	.33	[.26,.40]	.35	[.28,.41]	.19	[.10,.29]
ASD2	.38	[.34,.42]	.30	[.24,.37]	.35	[.29,.41]	.13	[.04,.23]
ASD3	.44	[.41,.48]	.34	[.27,.41]	.38	[.32,.44]	.19	[.09,.28]
MAS	.46	[.42,.50]	.36	[.29,.42]	.40	[.34,.46]	.25	[.15,.34]

Notes. CBCL= Parent reports (6;0 to 18;11 years); YSR= Self reports (11;0 to 18;11 years); TRF= Teacher reports (6;0 to 18;11 years); *r* = Correlation after Pearson; *Cl*= 95% Confidence interval; SPS = Stress Problems scale (Achenbach & Rescorla, 2007); OCP = Obsessive-Compulsive Problems scale (Achenbach & Rescorla, 2007); SCT = Sluggish Cognitive Tempo scale (Achenbach & Rescorla, 2007); <sup>1</sup>scale not available in YSR/11-18; DRS1 = Dysregulation scale (Ayer et al., 2009); DRS2 = Weighted and rounded Dysregulation scale (McQuillan, 2018); ASD1 = Autism Spectrum Disorder Scale (Ooi et al., 2011); ASD2 = Autism Spectrum Disorder Scale (So et al., 2012); ASD3 = Autism Spectrum Disorder Scale (Offermans et al., 2022); MAS = Mania Scale (Papachristou et al., 2013); \**n.s.*; all other correlations  $p \leq .01$

in the YSR. Most correlations with the DSM-oriented scales were at least moderate ( $r \geq .30$ ). All three ASD scales correlated significantly more strongly with INT than with EXT. With the exception of the Hyperactivity-Impulsivity Problems subscale, low to moderate correlations emerged for all three ASD scales with the DSM-oriented scales across all samples and informants. The final supplementary scale considered, MAS, showed higher correlations with EXT than with INT across informants and samples. Overall, correlations with INT as well as EXT were above .50 in the CBCL (apart from INT in the clinical CBCL sample) and the YSR, while in the TRF sample, only EXT reached a high coefficient ( $r = .89$ ). Apart from Somatic Problems in the clinical CBCL sample, all DSM-oriented scales correlated at least moderately with MAS in the CBCL and YSR samples. Furthermore, most DSM-oriented scales showed high correlations in the clinical TRF sample, although correlations with Depressive Problems and Anxiety Problems were low and the correlation with Somatic Problems did not even reach statistical significance.

#### **Mean differences between clinical and community sample**

Table 4 presents the mean differences between the clinical and community samples for the supplementary scales of the CBCL and YSR. As expected, scale scores were significantly higher in the clinical samples than in the community samples. Parent reports consistently showed larger effect sizes ( $1.00 \leq d \leq 1.75$ ). Effect sizes were somewhat smaller for the self-report, but nevertheless still categorized as large ( $0.85 \leq d \leq 1.03$ ), with the exception of PQS.

#### **Comparison between clinical samples with and without nearest respective diagnosis**

Table 5 presents the mean differences of supplementary scales between subgroups of the clinical sample with and without the nearest respective diagnosis i.e., diagnosis of obsessive-compulsive disorder for OCP and diagnosis of pervasive developmental disorders for ASD1, ASD2, and ASD3. Most differences were statistically significant at  $p \leq .001$ . As expected, the subgroup with the nearest respective diagnosis showed higher scale scores compared to the remaining clinical sample. Large effect sizes ( $d \geq 1.23$ ) were noted for the CBCL scales. In the YSR, only OCP differed between the two subgroups, with a large effect. Furthermore, in the TRF, differences concerning all ASD scales reached large effects, while the difference in the OCP was only moderate.

#### **Discussion**

The present study analyzed the supplementary scales of the German school-age versions of the Achenbach questionnaires regarding internal consistency and aspects of validity across informants and samples. The discussion focuses first on the scales that showed at least adequate internal consistency (SPS, DRS, MAS, PQS, ASD3) and then on the scales which do not show this cross-informantly. As our community sample was part of the multicultural analysis by Achenbach & Rescorla, our findings in the CBCL and YSR show similarity to the authors' findings concerning the 2007 scales (SPS, OCP, SCT, PQS) (9).

The internal consistency of the Stress Problems Scale (SPS) across CBCL and YSR and samples of different origin was at least adequate ( $\Omega > .70$ ) with a narrow exception in TRF ( $\Omega = .69$ ). This is in line with

**TABLE 4.** Mean differences between clinic and community samples for the CBCL/6-18 and YSR/11-18.

Origin N	CBCL/6-18					YSR/11-18							
	Clinic 1288		Community 2471		d	Clinic 755		Community 1797 (n <sub>PQS</sub> = 1791)		t	df	d	
Scale	M	SD	M	SD	t	df	M	SD	M	SD	t	df	d
SPS	9.07	4.56	2.86	2.90	44.42	1843.22	8.89	5.02	4.57	3.81	21.24	1135.13	1.03
OCP	3.18	2.86	0.80	1.27	28.43	1558.42	3.92	3.37	1.56	1.96	17.96	974.37	0.96
SCT <sup>1</sup>	1.79	1.77	0.51	0.91	24.25	1649.10							
PQS <sup>2</sup>							18.83	4.37	17.09	6.19	8.02	1974.84	0.30
DRS1	24.32	12.22	8.32	7.57	42.92	1816.33	21.33	11.34	12.71	9.51	18.35	1220.90	0.85
DRS2	24.30	11.98	8.23	7.54	43.80	1831.65	22.11	11.62	12.93	9.71	19.09	1217.43	0.89
ASD1	3.56	2.81	0.98	1.42	30.99	1640.01	3.57	2.70	1.65	1.89	17.77	1077.19	0.89
ASD2	3.99	3.13	1.08	1.59	31.34	1638.54	3.97	2.88	1.82	1.99	18.74	1069.62	0.94
ASD3	6.76	4.87	1.97	2.63	32.84	1686.72	6.69	4.37	3.40	3.30	18.59	1131.57	0.90
MAS	8.26	5.25	2.73	3.04	34.83	1748.99	7.94	4.78	4.93	4.13	15.09	1249.71	0.69

Notes: CBCL/6-18 = Parent reports (children and youths 6;0 to 18;11 years); YSR/11-18 = Self reports (11;0 to 18;11 years); TRF/6-18 = Teacher reports (children and youths 6;0 to 18;11 years); nPQS = Community sample size Positive Qualities; t = t value based on Welch-Test in service of uniformity for all scales as majority showed heterogeneity of variance between both samples; df = Degrees of freedom; d = Cohen's effect size; SPS = Stress Problems scale (Achenbach, 2007); OCP = Obsessive-Compulsive Problems scale (Achenbach, 2007); SCT = Sluggish Cognitive Tempo scale (Achenbach, 2007); 1scale not available in YSR/11-18; PQS = Positive Qualities scale (Achenbach, 2007); 2scale not available in CBCL/6-18 and TRF/6-18; DRS1 = Dysregulation scale (Ayer, 2009); DRS2 = Weighted and rounded Dysregulation scale (McQuillan, 2018); ASD1 = Autism Spectrum Disorder Scale (Ooi et al., 2011); ASD2 = Autism Spectrum Disorder Scale (So et al., 2012); ASD3 = Autism Spectrum Disorder Scale (Offermans et al., 2022); MAS = Mania Scale (Papachristou et al., 2013); All group differences are statistically significant p<.001

TABLE 5. Mean differences of supplementary scales between clinic subgroups with and without the nearest respective diagnosis<sup>1</sup>

Scale	Group	CBCL/6-18						YSR/11-18						TRF/6-18								
		n	M	SD	t	df	p	d	n	M	SD	t	df	p	d	n	M	SD	t	df	p	d
OCP	Y	74	6.85	3.37	-	78.70	<.001	1.44	62	6.95	3.62	-	753	<.001	1.02	32	3.25	2.27	-3.52	834	<.001	0.64
	N	1214	2.96	2.67	9.77 <sup>w</sup>			693	3.65	3.22	7.67				804	1.76	2.34					
ASD1	Y	68	6.96	3.48	-	71.36	<.001	1.33	29	5.14	3.17	-	753	.001	0.61	53	5.57	3.01	-8.16	834	<.001	1.16
	N	1220	3.37	2.64	8.37 <sup>w</sup>			726	3.50	2.66	3.22				783	2.58	2.54					
ASD2	Y	68	7.79	3.60	-	1286	<.001	1.34	29	5.10	3.31	-	753	.030	0.41	53	7.43	4.05	-	56.74	<.001	1.18
	N	1220	3.78	2.96	10.74			726	3.92	2.85	2.17				783	3.49	3.29	6.93 <sup>w</sup>				
ASD3	Y	68	12.24	4.96	-9.87	1286	<.001	1.23	29	8.59	4.62	-	753	.017	0.45	53	10.25	5.58	-6.60	834	<.001	0.94
	N	1220	6.45	4.69				726	6.61	4.34	2.39				783	5.73	4.77					

Notes. <sup>1</sup> "Obsessive-compulsive disorder" for OCP and "Pervasive developmental disorders" for ASD1, ASD2 and ASD3; CBCL/6-18 = Parent reports (children and youths 6;0 to 18;11 years); YSR/11-18 = Self reports (11;0 to 18;11 years); TRF/6-18 = Teacher reports (children and youths 6;0 to 18;11 years); Y = 'Yes' clinic sample with respective diagnosis; N = 'No' clinic sample without the respective diagnosis; n = sample size; M = Mean scale score; SD = Standard Deviation; t = t value; <sup>w</sup> Welch-Test for heterogeneity of variance; df = degrees of freedom; **p** = two-tailed significance; **d** = Cohen's effect size; OCP = Obsessive-Compulsive Problems scale (Achenbach, 2007); ASD1 = Autism Spectrum Disorder Scale (Ooi et al., 2011); ASD2 = Autism Spectrum Disorder Scale (So et al., 2012); ASD3 = Autism Spectrum Disorder Scale (Offermans et al., 2022)

previous multicultural findings reporting Cronbach's  $\alpha$  in the community sample for all three informant types (9) as well as with the results of You et al. (25) regarding their TRF ( $\alpha=.79$ ) and the YSR ( $\alpha=.85$ ) community sample. Moreover, our results also revealed adequate internal consistencies in the clinical sample across parents and self reports, whereas Zandberg et al. (13) found a somewhat lower internal consistency ( $\alpha=.68$ ) in self-reports in a group of 14-19-year-old girls with a primary diagnosis of posttraumatic stress disorder. Cross-informant correlations in the community sample were higher than in the multicultural analysis by Achenbach & Rescorla (9), who reported low or no correlations between parents and teachers as well as between teachers and adolescents in a multicultural community sample (9), as was the case in our clinical sample. Correlations for INT and EXT were high across all informants, and the mean difference in the SPS between the clinical and community sample was the highest of all supplementary scales (dCBCL = 1.75; dYSR = 1.03).

Both versions of the Dysregulation Scale (DRS1 and DRS2) achieved good ( $\Omega \geq .80$ ) to excellent ( $\Omega \geq .90$ ) internal consistencies, which was consistent with previous findings in a community sample reported by McQuillan et al. (19). However, upon closer inspection, it emerged that some items, such as Fears, Fears school, or Fearful, did not reach  $r > .30$  across informants. Furthermore, the item-total correlation indicated higher values for the YSR and particularly for items describing covert behavior (Cries a lot, Dreams, Feels guilty, Talks about suicide, Worries). While Rescorla et al. found minor effects of age and gender in a CBCL and YSR community sample (36), we confirmed low effects (if at all) in the clinical samples. For the TRF, boys scored slightly higher than girls, which is in line with the US findings (23). Across the CBCL and YSR samples, high correlations with EXT as well as INT emerged, with higher values for EXT, whereas TRF correlations with INT were only in the medium range for clinical ratings. Cross-informant correlations were at least moderate, with the exception of the TRF with the YSR. Discriminant validity was demonstrated across informants. All in all, both DRS performed good concerning internal consistency and the aspects of validity analyzed, with only minor differences between the unweighted and the weighted, rounded scale. Nevertheless, in the interest of parsimony, a reduced scale would be advantageous, especially since a look at the item-related analyses support this hypothesis. One possible approach would be a data-driven procedure, comparable with the optimizing of the ASD3 (22). Another approach could be the strategy used by Evans et al. (38) analyzing a selected

theoretically item-set based on previous research by confirmatory factor analyses.

The internal consistencies of the Mania Scale (MAS) are at least adequate ( $\Omega \geq .80$ ). But our results were lower than in previous studies (23, 30), though we were able to confirm the internal consistencies for a broader age range and across genders. Across the informants and samples, high correlations with EXT were found, whereas high correlations with INT were only found in the community samples. Cross-informant correlations were at least moderate, except for the TRF with the YSR. Mean differences between the clinical and community sample were large for the CBCL and medium for the YSR.

In the community sample, the YSR-specific scale Positive Qualities Scale (PQS) showed good internal consistency, in line with the findings of Achenbach & Rescorla (9), the clinical sample it was adequate. The correlations with INT as well as EXT, as far as significant, were at least low, due to a low variance of PQS. While no linear correlation between PQS and INT was found in the clinical sample, positive correlations with both EXT and INT were found in the community sample. PQS may therefore be interpreted as an enriching concept because the positive scale stands out from all other - problem-oriented - scales. On the item level, a need for further research seems to be apparent, as some items had a rather low item-total correlation. In contrast to Rescorla et al. (16), PQS was only slightly, if at all, correlated with gender. Age effects within the community sample were small and in line with the literature (16). The effects were also small in our clinical sample, but in the opposite direction. The discriminant validity was rather low, with the clinical sample unexpectedly achieving slightly higher values. Of the three Autism Spectrum Disorder scales (ASD), Offermans et al.'s (22) data-driven ASD3 showed at least adequate internal consistencies across all samples and informants, supporting and extending the findings of the original research group. Although the ASD3 was originally developed based on parent reports, we found even better internal consistency for the TRF. The analyses of convergent/ divergent validity for all samples and perspectives concerning ASD resulted in high correlations between all three scales and INT. Additionally, ASD2 and ASD3 correlated highly with EXT, but correlations with INT were significantly higher in all cases. Moreover, cross-informant correlations were consistently moderate, with the exception of the TRF with the YSR. Discriminant validity was proven for all ASD scales, whereas convergent/ divergent validity was high for parent and teacher reports but at least low for self-reports. Gender and age effects were only slightly pronounced, if at all.

With regard to internal consistency, the Obsessive-Compulsive Problems scale (OCP) showed predominantly weak internal consistency in the community sample, which was in line with the findings of Achenbach & Rescorla (9) but contrary to those of Nelson et al. (14) and Geller et al. (26). In the present study, OCP only in the clinical YSR sample achieved an adequate internal consistency. Therefore, the further results should only be interpreted as meaningful for this informant perspective. The high correlation of OCP with INT was significantly higher than that with EXT, in accordance with clinical expectation. The cross-informant correlations between the CBCL and YSR in the clinical and community samples were at best moderate, although higher overall compared to Achenbach & Rescorla (9). This is unsurprising from a clinical perspective, as OCP symptoms are unlikely to be fully communicated by youth, and are more likely to be hidden. Discriminant validity was given across informants: Large mean differences were found between the respective nearest diagnosis group and the remaining clinical sample with the CBCL and YSR, and medium ones with the TRF. The correlation analyses of OCP with gender and age revealed only one moderate effect, which was a gender effect in the clinical YSR sample.

For the Sluggish Cognitive Tempo scale (SCT), internal consistencies were good in the clinical TRF sample ( $\Omega=.81$ ) but poor in both CBCL samples. This is in line with the results of Achenbach and Rescorla (8), while Bauermeister et al. reported good internal consistencies for their clinical CBCL and TRF samples (9, 27)). Furthermore, Achenbach and Rescorla (9) found adequate internal consistencies for TRF in a community sample (not available for our presented analyses). Since the SCT was developed based on the TRF, our results support the assumption that this scale might be more, if not exclusively, suitable for the TRF. As also reported by Bauermeister et al. (39), we detected strong correlations between SCT and INT in our clinical sample. Despite higher item overlaps between SCT and EXT, the correlations were low, and indeed significantly lower than those with INT and not even significant for the TRF (clinical sample). Becker et al. (28) likewise found no significant correlations of SCT with the DSM-oriented scales Attention Deficit/Hyperactivity Problems or Oppositional Defiant Problems in the CBCL. However, our low to moderate results should be interpreted with caution due to poor internal consistency. In contrast, in our TRF sample, SCT did not even correlate with any Hyperactivity-Impulsivity scale, with Aggressive Behavior, or with the DSM-oriented scale Oppositional Defiant Problems.

Surprisingly, the following scales correlated notably strongly with SCT across the samples: Withdrawn/depressed (one item in common), the DSM-oriented scale Depressive Problems (two items in common), as well as moderately with the DSM-oriented scale Inattention (no items in common) in the TRF. On the one hand, the strong correlation between SCT and Depressive Problems might imply that daydreaming from SCT is associated with rumination (28, 40); a moderate correlation with the dimension Inattention may thus also be rooted in this. On the other hand, the strong correlation between SCT and Withdrawn/depressed can be explained by the associated social withdrawal (15, 28, 40). In view of the lack of the internal consistency of the CBCL scale, the moderate correlation between the CBCL and TRF and the large effect of discriminant validity for the CBCL should be interpreted with caution.

### ***Limitations and Perspectives***

The following limitations should be noted when interpreting the results. In the present work, the items were analyzed in their original ordinal form, while some research groups followed the recommendations of Achenbach and Rescorla (2001) (3) to aggregate the items in a binary code, and others did not describe their precise evaluation strategy. These different approaches might affect the comparability of the results. Another reason for a lack of comparability may lie in the replacement of the items that were newly introduced in the 2001 version of the questionnaires with mean item scores of the respective traditional problem scale. Although missing data were handled based on the recommendations of Achenbach and Rescorla (3), referring to a maximum of 8 items missing out of the total score, which were set to 0, this might have concerned items of the supplementary scales. Given that the scales comprised differing numbers of items, it cannot be ruled out that there were cases in which more than 10% of the items in the respective scale fell under this criterion. As a further problem of our analysis, it should be considered that we used uncorrected correlations between scales that have common items. Clearly, this induced higher values of the correlations, but it seemed appropriate as we wished to analyze the closeness of concepts and the probable gain from the supplementary scales. Furthermore, it also seems to be of great interest for further research to elicit combined scales between informants, as has already been done for the ASD2 (21). Another interesting approach would be to examine the item level for the different informants (e.g., intraclass correlation). Exploring the construct validity of all of the supplementary scales, as previously done for the ASD3 (22), may be a further

relevant research endeavor. Reliability testing with more than cross-sectional data as well as validity testing with other multimodal and psychometric instruments might also be extremely helpful for clinical practice. We limit ourselves more to examining the gain from the additional scales compared to the already existing scales. The potential of the supplementary scales in the clinical TRF sample, as revealed in this study, supports the inclusion of a TRF community sample for future research on the supplementary scales, and underlines the importance of cross-informant diagnostic assessment. Overall, we recommend for further research to test the supplementary scales using the analysis strategies presented here to enable.

### Clinical Significance

The Mania Scale and the Autism Spectrum Disorder scale according to Offermans et al. (22) can be recommended for clinical use. Inspired by Offermans et al.'s successful strategy and in the interest of parsimony, instead of using the current versions of the Dysregulation Scale, we suggest a new approach of data-driven analyses with the aim of focusing on the most relevant items in terms of content. Construct and discriminant validity suggest the Stress Problems Scale as a higher-order dimension. If at all, the Obsessive-Compulsive Problems scale should only be used in self-assessment, while valid and reliable results on the Sluggish Cognitive Temposcale are more likely in teacher assessment. The Positive Qualities scale seems to be an enriching concept but needs further research.

### Ethics approval and consent to participate

The community sample originates from the nationally representative sample of the PAK-KID study, which was collected with the German version of the Child Behaviour Checklist and the Youth Self-Report by Achenbach. The data collection took place within the framework of a multi-topic survey in representatively selected private households, completely anonymous to us. The sample has already been used for example to generate the German norms for the CBCL and YSR (1). For youths under the age of 16 years (minors) informed consent to participate was obtained from the parents or legal guardians. Data collection has been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All persons gave their informed consent prior to their inclusion in the study. Clinical data collection for this study has been approved by the ethics committee at the University of Cologne (15-401) and has, therefore, been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All persons gave their informed consent prior to their inclusion in the study. For youths under the age of 16 years

(minors) informed consent to participate was obtained from the parents or legal guardians.

### Availability of data and materials

The pseudonymized datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

### Declaration of interests

As part of the Working Group German Child Behavior Checklist, MD and JP are authors of the German-language ASEBA questionnaire versions for school age and the corresponding manual, and therefore receive royalties from the publisher Hogrefe. LN and EK declare that they are not aware of any conflicts of interest.

### Funding

This work was not supported by special funding but the authors are staff scientists at the mentioned affiliative institutions.

### Acknowledgements

We thank the patients and their families and the staff at the School for Child and Adolescent Cognitive Behavior Therapy (AKiP), Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany, who are the source of the clinical dataset. We are grateful to the working group German Child Behavior Checklist for providing the community sample used in these analyses.

### References

1. Döpfner M, Plück J, Kinnen C. CBCL/6-18R, TRF/6-18R, YSR/11-18R. Deutsche Schulalter-Formen der Child Behavior Checklist von Thomas M. Achenbach. Göttingen: Hogrefe; 2014.
2. Döpfner, M., Görtz-Dorten, A., & Petermann, F. (2024). Diagnostik psychischer Störungen im Kindes- und Jugendalter (Vol. 2). Hogrefe Verlag GmbH & Company KG.
3. Achenbach TM, Rescorla L. Manual for the ASEBA school-age forms & profiles: An integrated system of multi-informant assessment. Burlington, Verm.: ASEBA; 2001.
4. Ivanova MY, Achenbach TM, Rescorla LA, Guo J, Althoff RR, Kan K-J, et al. Testing Syndromes of Psychopathology in Parent and Youth Ratings Across Societies. *J Clin Child Adolesc Psychol*. 2019;48:596–609. doi:10.1080/15374416.2017.1405352.
5. Ivanova MY, Dobrea A, Döpfner M, Erol N, Fombonne E, Fonseca AC, et al. Testing the 8-syndrome structure of the child behavior checklist in 30 societies. *J Clin Child Adolesc Psychol*. 2007;36:405–17. doi:10.1080/15374410701444363.
6. Ivanova MY, Achenbach TM, Rescorla LA, Dumenci L, Almqvist F, Bathiche M, et al. Testing the Teacher's Report Form Syndromes in 20 Societies. 2007;36:468–83.
7. Ivanova MY, Achenbach TM, Rescorla LA, Dumenci L, Almqvist F, Bilenberg N, et al. The generalizability of the Youth Self-Report syndrome structure in 23 societies. *J Consult Clin Psychol*. 2007;75:729–38. doi:10.1037/0022-006X.75.5.729.
8. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5-TR™. Washington, DC: American Psychiatric Association Publishing; 2022.

9. Achenbach TM, Rescorla L. Multicultural Supplement to the Manual for the ASEBA School-Age Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families; 2007.
10. Wolfe VV, Gentile C, Wolfe DA. The impact of sexual abuse on children: A PTSD formulation. *Behavior Therapy*. 1989;20:215–28. doi:10.1016/S0005-7894(89)80070-X.
11. Ruggiero KJ, McLeer SV. PTSD scale of the Child Behavior Checklist: concurrent and discriminant validity with non-clinic-referred sexually abused children. *J Trauma Stress*. 2000;13:287–99. doi:10.1023/A:1007710828777.
12. Achenbach TM, Rescorla LA, Ivanova MY. Guide to family assessment using the ASEBA. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families; 2015.
13. Zandberg L, Kaczurkin AN, McLean CP, Rescorla L, Yadin E, Foa EB. Treatment of Adolescent PTSD: The Impact of Prolonged Exposure Versus Client-Centered Therapy on Co-Occurring Emotional and Behavioral Problems. *J Trauma Stress*. 2016;29:507–14. doi:10.1002/jts.22138.
14. Nelson EC, Hanna GL, Hudziak JJ, Botteron KN, Heath AC, Todd RD. Obsessive-compulsive scale of the child behavior checklist: specificity, sensitivity, and predictive power. *Pediatrics*. 2001;108:E14. doi:10.1542/peds.108.1.e14.
15. Carlson CL, Mann M. Sluggish cognitive tempo predicts a different pattern of impairment in the attention deficit hyperactivity disorder, predominantly inattentive type. *J Child Adolesc Psychol*. 2002;31:123–9. doi:10.1207/S15374424JCCP3101\_14.
16. Rescorla L, Achenbach TM, Ivanova MY, Dumenci L, Almqvist F, Bilenberg N, et al. Epidemiological comparisons of problems and positive qualities reported by adolescents in 24 countries. *J Consult Clin Psychol*. 2007;75:351–8. doi:10.1037/0022-006X.75.2.351.
17. Ayer L, Althoff R, Ivanova M, Rettew D, Waxler E, Sulman J, Hudziak J. Child Behavior Checklist Juvenile Bipolar Disorder (CBCL-JBD) and CBCL Posttraumatic Stress Problems (CBCL-PTSP) scales are measures of a single dysregulatory syndrome. *J Child Psychol Psychiatry*. 2009;50:1291–300. doi:10.1111/j.1469-7610.2009.02089.x.
18. Biederman J, Wozniak J, Kiely K, Ablon S, Faraone S, Mick E, et al. CBCL Clinical Scales Discriminate Prepubertal Children with Structured Interview—Derived Diagnosis of Mania from Those with ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry*. 1995;34:464–71.
19. McQuillan ME, Kultur EC, Bates JE, O'Reilly LM, Dodge KA, Lansford JE, Pettit GS. Dysregulation in children: Origins and implications from age 5 to age 28. *Dev Psychopathol*. 2018;30:695–713. doi:10.1017/S0954579417001572.
20. Ooi YP, Rescorla L, Ang RP, Woo B, Fung DSS. Identification of autism spectrum disorders using the Child Behavior Checklist in Singapore. *J Autism Dev Disord*. 2011;41:1147–56. doi:10.1007/s10803-010-1015-x.
21. So P, Greaves-Lord K, van der Ende J, Verhulst FC, Rescorla L, Nijs, FA de Nijs, Pieter. Using the Child Behavior Checklist and the Teacher's Report Form for identification of children with autism spectrum disorders. *Autism*. 2013;17:595–607. doi:10.1177/1362361312448855.
22. Offermans JE, Bruin EI de, Lange AMC, Middeldorp CM, Wesseldijk LW, Boomsma DI, et al. The Development and Validation of a Subscale for the School-Age Child Behavior Checklist to Screen for Autism Spectrum Disorder. *J Autism Dev Disord*. 2022. doi:10.1007/s10803-022-05465-7.
23. Papachristou E, Ormel J, Oldehinkel AJ, Kyriakopoulos M, Reinares M, Reichenberg A, Frangou S. Child Behavior Checklist-Mania Scale (CBCL-MS): development and evaluation of a population-based screening scale for bipolar disorder. *PLoS One*. 2013;8:e69459. doi:10.1371/journal.pone.0069459.
24. Rescorla LA, Blumenfeld MC, Ivanova MY, Achenbach TM, International ASEBA Consortium. International Comparisons of the Dysregulation Profile Based on Reports by Parents, Adolescents, and Teachers. *J Clin Child Adolesc Psychol*. 2019;48:866–80. doi:10.1080/15374416.2018.1469090.
25. You DS, Youngstrom EA, Feeny NC, Youngstrom JK, Findling RL. Comparing the Diagnostic Accuracy of Five Instruments for Detecting Posttraumatic Stress Disorder in Youth. *J Clin Child Adolesc Psychol*. 2017;46:511–22. doi:10.1080/15374416.2015.1030754.
26. Geller DA, Doyle R, Shaw D, Mullin B, Coffey B, Petty C, et al. A quick and reliable screening measure for OCD in youth: reliability and validity of the obsessive compulsive scale of the Child Behavior Checklist. *Compr Psychiatry*. 2006;47:234–40. doi:10.1016/j.comppsy.2005.08.005.
27. Bauermeister JJ, Matos M, Reina G, Salas CC, Martínez JV, Cumba E, Barkley RA. Comparison of the DSM-IV combined and inattentive types of ADHD in a school-based sample of Latino/Hispanic children. *J Child Psychol Psychiatry*. 2005;46:166–79. doi:10.1111/j.1469-7610.2004.00343.x.
28. Becker SP, Withrow AR, Stoppelbein L, Luebke AM, Fite PJ, Greening L. Sluggish cognitive tempo is associated with suicide risk in psychiatrically hospitalized children. *J Child Psychol Psychiatry*. 2016;57:1390–9. doi:10.1111/jcpp.12580.
29. Raiker JS, Greening L, Stoppelbein L, Becker SP, Fite PJ, Luebke AM. Mediating Effect of Psychopathy on the Risk of Social Problems Among Children with ADHD Versus Sluggish Cognitive Tempo Symptoms. *Child Psychiatry Hum Dev*. 2015;46:523–32. doi:10.1007/s10578-014-0493-z.
30. Papachristou E, Schulz K, Newcorn J, Bédard A-CV, Halperin JM, Frangou S. Comparative Evaluation of Child Behavior Checklist-Derived Scales in Children Clinically Referred for Emotional and Behavioral Dysregulation. *Front Psychiatry*. 2016;7:146. doi:10.3389/fpsy.2016.00146.
31. Evers A, Hagemester C, Hostmølingen A, Lindley P, Muñoz J, Sjöberg A. EFPA review model for the description and evaluation of psychological tests: test review form and notes for reviewers. 2013. <https://psyktestbarn.r-bup.no/cms/files/913>. Accessed 10 Oct 2023.
32. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Erlbaum; 1988.
33. Hemmerich WA. *StatistikGuru: Korrelationen statistisch vergleichen*. 2017. <https://statistikguru.de/rechner/korrelationen-vergleichen.html>. Accessed 23 Mar 2023.
34. Hittner JB, May K. DEPCORR: A SAS Program for Comparing Dependent Correlations. *Applied Psychological Measurement*. 1998;22:93–4. doi:10.1177/01466216980221010.
35. Dunn OJ, Clark V. Correlation Coefficients Measured on the Same Individuals. *Journal of the American Statistical Association*. 1969;64:366. doi:10.2307/2283746.
36. Hittner JB, May K, Silver NC. A Monte Carlo evaluation of tests for comparing dependent correlations. *J Gen Psychol*. 2003;130:149–68. doi:10.1080/00221300309601282.
37. Rescorla LA, Blumenfeld MC, Ivanova MY, Achenbach TM, International AC. International Comparisons of the Dysregulation

Profile Based on Reports by Parents, Adolescents, and Teachers. *J Clin Child Adolesc Psychol.* 2019;48:866–80. doi:10.1080/15374416.2018.1469090.

38. Evans SC, Bonadio FT, Bearman SK, Ugueto AM, Chorpita BF, Weisz JR. Assessing the Irritable and Defiant Dimensions of Youth Oppositional Behavior Using CBCL and YSR Items. *J Clin Child Adolesc Psychol.* 2020;49:804–19. doi:10.1080/15374416.2019.1622119.
39. Bauermeister JJ, Barkley RA, Bauermeister JA, Martínez JV, McBurnett K. Validity of the sluggish cognitive tempo, inattention, and hyperactivity symptom dimensions: neuropsychological and psychosocial correlates. *J Abnorm Child Psychol.* 2012;40:683–97. doi:10.1007/s10802-011-9602-7.
40. Becker SP, Webb KL, Dvorsky MR. Initial Examination of the Bidirectional Associations between Sluggish Cognitive Tempo and Internalizing Symptoms in Children. *J Clin Child Adolesc Psychol.* 2021;50:258–66. doi: 10.1080/15374416.2019.1630836.