NEW RESEARCH AGGRESSIVE BEHAVIOR

Check for updates

Psychometric Features, Score Distributions, and Factor Structure of the Retrospective Modified Overt Aggression Scale From a Pediatric Cohort Referred for Behavioral Health Treatment

Joseph C. Blader, PhD

Objective: Aggression is a complication of many psychiatric conditions in youth, but a need remains to measure its specific behaviors. This study evaluated the psychometric and other features of the Retrospective-Modified Overt Aggression Scale (R-MOAS), a 16-item, adult-informant measure for the frequency of verbal, property-related, physical, and self-directed aggressive behaviors.

Method: Parents of 4,155 youth, aged 5 to 17 years, completed the R-MOAS following referral for behavioral health concerns from general pediatric settings. Analyses examined the following: (1) score distributions, (2) internal consistency, test–retest reliability, and validity, (3) item response theory (IRT) performance, and (4) factor structure.

Results: Scores best fit a zero-modified exponential distribution. Self-directed aggressive behavior decreased less with age among female patients. Cronbach α and McDonald ω were high (0.88 and 0.87, respectively), indicating good internal consistency. Test–retest reliability was 0.70. The pattern of correlations with other measures demonstrated convergent and discriminant validity. IRT analyses showed good discrimination covering a range of scores. IRT supports the ordinality of ratings within items but not the scale's traditional approach to weighting item severity. Factor analysis suggested a 2-factor structure. One factor has high loadings from verbal items and milder physical and property-directed aggression ("Eruptive"), and the other factor's loadings drew from self-directed and more destructive behaviors ("Harmful/Distressed"). Measures of affective disturbances made unique contributions to the Harmful/Distressed factor only, whereas the Eruptive factor showed stronger influences of impulsiveness and externalizing behavior.

Conclusion: The R-MOAS fulfills psychometric criteria for reliability, validity, and IRT performance. It can be a useful component in clinical care and research for the identification, quantification, and outcome monitoring of aggressive behavior in youth. Scoring using item scores is superior to the weighting methods of prior versions, which should be disfavored in youth populations. Factor structure suggests one phenotype that features verbal and relatively minor forms of aggression and another in which self-directed and severe harmful behaviors accompany greater affective disturbance.

Plain language summary: Aggressive behavior is a common and serious concern among youth receiving mental health care. Measuring this behavior is important for clinical care and research. This study shows that a parent-report rating scale, the Revised – Modified Overt Aggression Scale (R-MOAS), is useful based on analysis of scales completed for over 4,000 children and adolescents in pediatric/psychiatric collaborative care settings. Its psychometric properties fulfill standard criteria for reliability and validity. The article also presents the frequencies of several types of aggressive behavior in this sample and demonstrates that these behaviors are frequent and often severe. For example, almost a quarter were reported to have struck another person in the prior week once or twice and another 12% were said to have done so 3 or 4 times. Because there are few well-validated scales to assess aggressive behavior clinically, this study supports the use of the R-MOAS to fill this gap.

Key words: behavior rating scale; aggression; psychometrics; child behavior; adolescent behavior

JAACAP Open 2025;3(1):85-100.

Aggressive behavior is a common and often severe complication of numerous psychiatric disturbances among young people, and its quantification is important to clinical assessment, outcome monitoring, and research. First, harmful behavior is among the leading reasons for children and adolescents to be psychiatrically hospitalized and to present at crisis services.[1-3] Measurement of these urgent concerns beyond clinical narrative improves clinical triage, disposition, and communication. Second, aggression is the symptom for which children without psychotic illness most often receive treatment with antipsychotic medications and multi-agent regimens[4-6]; the adverse effect risks of these interventions warrant unambiguous measurement of their indications for use and to determine their efficacy. Third, aggressive behavior has a corrosive impact on relationships and contributes to social marginalization that further worsens the outlook for achieving age-normative role functioning.[7-9] Supports for affected youngsters therefore require outcomes assessments that include the frequency and severity of specific aggressive behaviors. Fourth, growing research interest in the transdiagnostic nature of many behavioral disturbances, including aggression, requires tools to evaluate behaviors distinctly to determine their relationships with cross-cutting higher-order constructs (eg, anger, hostility, irritability, lability, impulsivity, inflexibility).[10]

Definitions of aggression differ, but, at a behavioral level, generally converge on describing nonaccidental action that is harmful or hurtful, that violates social norms, and that may include both physical and verbal behavior.[11,12] In psychiatric contexts, self-harming is often included, and the empirical association between interpersonal and self-directed aggression[13,14] supports the inclusion of both in the same assessment framework.[15,16] The specific behaviors a clinical rating scale for aggression might encompass therefore include verbal lashing out, interpersonal physical aggression, damage to property, and harm directed toward oneself.

Several rating scales for youth include some aggressive behavior items but often comingle them in subscales with other externalizing behavior problems, such as defiance, hostility, argumentativeness, and poor rule adherence, which limits their precision as gauges of aggression itself.[10,17-21] A few scales focus on differentiating reactive (impulsive, frustrative, defensive) aggressive behavior from proactive (deliberate, extractive, volitional) aggression. However, they do not assess the absolute frequency and severity of these behaviors,[22] rely on youth self-report,[23] or include items that are relevant to the reactive/proactive construct but that are not aggressive behaviors per se.[22,23] The Children's Aggression Scale[24] is a 33-item parent-rated measure rating scale for children that queries aggressive acts toward specific targets (adults and other children, in and out of the home) and distinguishes provoked and unprovoked hitting. However, it lacks self-aggression items, and it enumerates forms of aggressive behavior with less specificity than other measures. A recent review of measures for reactive aggression notes the need for narrowly focused but comprehensive assessment of aggressive behavior, while observing that there are none with large-sample based norms and robust evaluation of reliability and validity.[21]

A few measures that quantify aggressive behaviors specifically were developed for adults. The Overt Aggression Scale (OAS) by Yudofsky *et al.*[16] was designed for use on inpatient psychiatric units. Specific aggressive behaviors were organized in 4 categories: verbal, physical toward others, physical toward oneself, and physical toward objects. Nursing staff used the form to record these behaviors as they occurred. The Modified Overt Aggression Scale (MOAS) by Kay *et al.*[25] is a revision that records only the most severe behavior in each category per episode and uses gradations in item and category severity to provide a weighted scoring system. The Overt Aggression Scale–Modified (OAS-M) by Coccaro *et al.*[15] is a further adaptation that uses an interview format in which adult outpatients report the number of temper outbursts during the preceding week, which are then categorized to the best-matching item.

The need for a parent- and teacher-completed measure of children's aggressive behavior for a stepped-treatment medication trial[26] led to piloting of the OAS-M items, with alterations in wording to improve relevance to children. This piloting revealed that scores using parents' reports of the number of times that behaviors occurred were not reliable. The adoption of Likert scaling for behavior frequencies resulted in better correspondence of scores with other clinical information obtained during study visits in 2 trials.[26,27] This instrument was called the Retrospective – Modified Overt Aggression Scale (R-MOAS).

This paper presents the score distribution, reliability, validity, item response theory (IRT) performance, and factor structure of the R-MOAS based on data from a collaborative-care pediatric/behavioral health service. It aims to fill the gaps in clinical measurement noted earlier and identified in recent reviews.[10,21] Although data from clinical samples do not provide population-level norms, they are useful to evaluate the properties of measures for behaviors that occur infrequently in general population samples. This study also compares 2 scoring approaches: (1) the MOAS and OAS-M convention of multiplying each item's score by the severity of the item's behavior to give greater weight in the total to more harmful or disruptive forms of aggression, vs (2) simply adding the item scores without severity weighting. The former has intuitive appeal but has not yet been evaluated relative to the latter simpler scoring method, and it is based on assumptions about item severity that require validation.

METHOD

Data Source

Data for this report came from the intake assessments of 4,155 individuals 5 to 17 years of age who were referred for a range of concerns to behavioral health clinicians embedded within 12 primary care pediatric settings. Sex recorded in medical records was male for 63% and female for 37%. Of the participants, 9.2% self-identified as Black or African American, 86.9% as White, 74% as Latin American, and 5% as another heritage.

Measures

Retrospective Modified Overt Aggression Scale. The R-MOAS contains 16 items divided evenly among 4 categories of behavior: Verbal, Physical, Property-Related, and Self-Directed. Supplemental Material, available online, contains the scale and scoring guide in the "Forms" section. Each item describes a specific behavior, and informants rate its frequency from 4 choices (3 in the Verbal section) arranged in increasing frequency to form a Likert-type scale. The 4 items in each category are also arranged in presumed order of severity. For example, the first item in the Physical category is “How many times did your child act like he/she was about to hit somebody or took a swing at someone without actually hitting another person?” The second item is “How many times did your child hit someone with hands or an object, kick, push, scratch or pull hair, without causing real injury?” Items 3 and 4 ask about aggressive behavior causing minor and severe injury, respectively. As with the MOAS and OAS-M, each item’s frequency score is multiplied with weights to accord larger values to more severe aggressive behaviors. The weighting system first multiplies the item score (based on frequency) by the item’s severity within its category (eg, the second item in each category is multiplied by 2, the third by 3, etc). Then, a second weight is applied based on the item’s category (Verbal = 1, Property = 2, Self = 3, Physical = 4). For example, if an informant selected the fourth choice (“5 or more times,” scored as 3) for Item 2 (weight of 2) in the Property-Related category (also weighted 2), the score for that item would be 12 (3 × 2 × 2). Analyses in this report that involve unweighted scores reflect the item’s frequency-based values without multiplication by weights.

Conners Global Index–Parent Form. The Conners Global Index–Parent Form (ConnGI)[28] is a widely used 10-item scale, in which 7 items constitute a Restless/Inattentive subscale and 3 items form an Emotional Lability subscale. Raw scores are converted to T scores using age- and sex-based norms for individuals 6 to 18 years of age.

Child Behavior Checklist. The Child Behavior Checklist (CBCL)[17] is a well-established parent-completed measure of child behaviors relevant to psychopathology. Its 112 items are organized into empirically derived clusters that make up its 8 subscales and its broad-band Internalizing and Externalizing components (each comprising 2 of the individual subscales). Raw scores are converted to T scores using age- and sex-based norms for youth 6 to 18 years of age.

Demographic Information. Age and sex information were recorded by clinical staff from medical records, and child’s race/ethnicity information was reported by caregivers.

Data Acquisition

Patients were referred by their primary care pediatrician to the behavioral health clinician, who was integrated into the primary care service, for assessment and treatment of behavioral concerns. The initial visit focused on comprehensive assessment and included completion of the R-MOAS, ConnGI-P, and CBCL by the youth’s caregiver. Subsequent office visits usually included re-administration of the R-MOAS and ConnGI, which enabled this study to estimate test–retest reliability for a subset of youth who were seen again within 30 days.

Data Analysis

Analyses were conducted with SAS 9.4,[29] and, for some of the IRT analyses, flexMIRT 3.6.[30] Given the small number of patients with missing R-MOAS items (n = 125, 3%), list-wise deletion, rather than imputation, was used for these data, yielding an effective sample of 4,030. To compare the influence of R-MOAS scoring methods on psychometric features, both unweighted scores (reflecting frequency only) and weighted scores were examined.

Analysis of score distributions was performed with SAS PROC UNIVARIATE and assessed by quantile–quantile plots. Score percentiles stratified by sex and year-of-age were calculated and plotted based on observed scores. The associations of scores with sex and age were tested in leastsquares regression models with SAS PROC GLM.

Scale internal consistency was measured via Cronbach’s α using SAS PROC CORR. An alternative index, with the more realistic assumption that item variances are not all equal, is McDonald’s ω, which uses the same metric (0 to 1).[31] McDonald’s ω was computed using the SAS macro by Hayes.[32] Test–retest reliability was estimated by the Spearman rank-order correlation (ρ) of intake evaluation scores with those obtained for patients whose next office visit was within 30 days.

IRT performance was evaluated using unidimensional and multidimensional graded-response models[33,34] with SAS PROC IRT and with flexMIRT.[30] Analyses evaluated the ordinality of choices *within* items, as well as the assumptions of increasing severity between items and between the behavioral categories upon which the weighted scoring approach is based.

Factor structure was examined via exploratory factor analysis (EFA) with half of the sample (n = 2,015), estimated using principal axis factoring and rotated using the oblique varimax method performed with PROC FACTOR. The number of factors was chosen based on scree plot and eigenvalue inspection, as well as via parallel analysis,[35] using the SAS macro %parallel.[36] Confirmatory factor analysis (CFA), using the other half (n = 2,015) as a validation sample, compared the EFA-generated solution with 2 *a priori* comparator models: (1) a single-factor model, and (2) a 4-factor model that treated the behavioral categories (Verbal, Physical, etc) as factors. Weighted least squares were used for estimation in CFA. Only unweighted R-MOAS scores were used in these analyses.

Validity was evaluated through R-MOAS scores’ Spearman correlations with total and subscale T scores of the ConnGI and CBCL. Participants 5 years of age were omitted from the validity analyses because the ConnGI and CBCL are not normed for those under 6 years. Factor scores derived from CFA were also examined for their bivariate (tested with Spearman ρ) and multivariable (via SAS PROC GLM) associations with these measures.

RESULTS

Distribution

Figure S1, available online, displays the complete distribution of R-MOAS Total scores for weighted and unweighted scores. In all, 901 (22%) had scores of zero, with a precipitous drop in the frequencies of higher scores. This discontinuity indicates that a zero-modified distribution, which censors the zero values, would be appropriate for estimating the best parametric cumulative density distribution. The histogram in Figure 1 shows the distribution of the zero-modified R-MOAS total scores and their fit with 3 parametric distributions: normal, lognormal, and exponential. The insets in Figure 1 contain their parameter estimates and assessment of fit. The exponential distribution had the best fit (lowest Cramér-von Mises ω^2), although fit for weighted scores ($\omega^2 = 0.52$) is better than for unweighted scores ($\omega^2 = 4.52$). The quantile–quantile (Q-Q) plots in Figure S2, available online, show R-MOAS observed total scores for the sample vs expected percentiles from the estimated exponential distribution parameters. For both weighted and unweighted scoring, observed values and values derived from estimated distribution parameters were tightly concordant until high scores (R-MOAS Total >97th percentile), although the association is tighter for weighted vs unweighted scoring.

Age and Sex Trends

Figure 2 shows the complete sample's median value and cut points for the 75th, 90th, and 95th percentiles of R-MOAS Total scores, weighted and unweighted, stratified by sex and age, along with means and SDs. (Figure S3, available online, contains this information for subscales. Table S1, available online, contains all response frequencies by sex and age.) For R-MOAS Total scores, there was a main effect in which values decrease with age (B $= -1.69$, $F_{1,4026} = 125.73$, $p < .0001$). There was no significant effect for sex (B $= -1.05$, $F_{1,4026} = 2.48$, $p = .12$). Looking at subscales, for Self-Directed aggression the age-by-sex interaction is significant for weighted ($F_{1,4026} = 4.22$, $p = .04$) and for unweighted ($F_{1,4026} = 3.6$, $p = .05$) scores, indicating a smaller decrease with age among female relative to male participants (Figure S3d and S3h).

Reliability

Table 1 contains reliability statistics for both weighted and unweighted scores. Unweighted scores had higher values for Cronbach's α, the traditional test for internal consistency reliability. This likely results from the weighted scores having higher inter-item variance. In both scoring approaches, all but the Self-Directed subscale were over 0.71 with the values for the Total Score of 0.84 (weighted) and 0.88 (unweighted).

The large number of observations (22%) with R-MOAS Totals of zero might inflate reliability estimates. Re-computation with a sample that reduced to 10% the proportion with a Total score of zero produced only slightly smaller but still high reliability values, with Cronbach's $\alpha = 0.83$ and 0.87.

Computed using Hayes' macro for SAS,[32] ω was 0.87 for weighted scores and 0.90 for unweighted with the full sample. In the zero-score reduced sample, ω was 0.85 (weighted) and 0.89 (unweighted). Omega (ω) values for subscales also showed higher reliabilities for the unweighted scores.

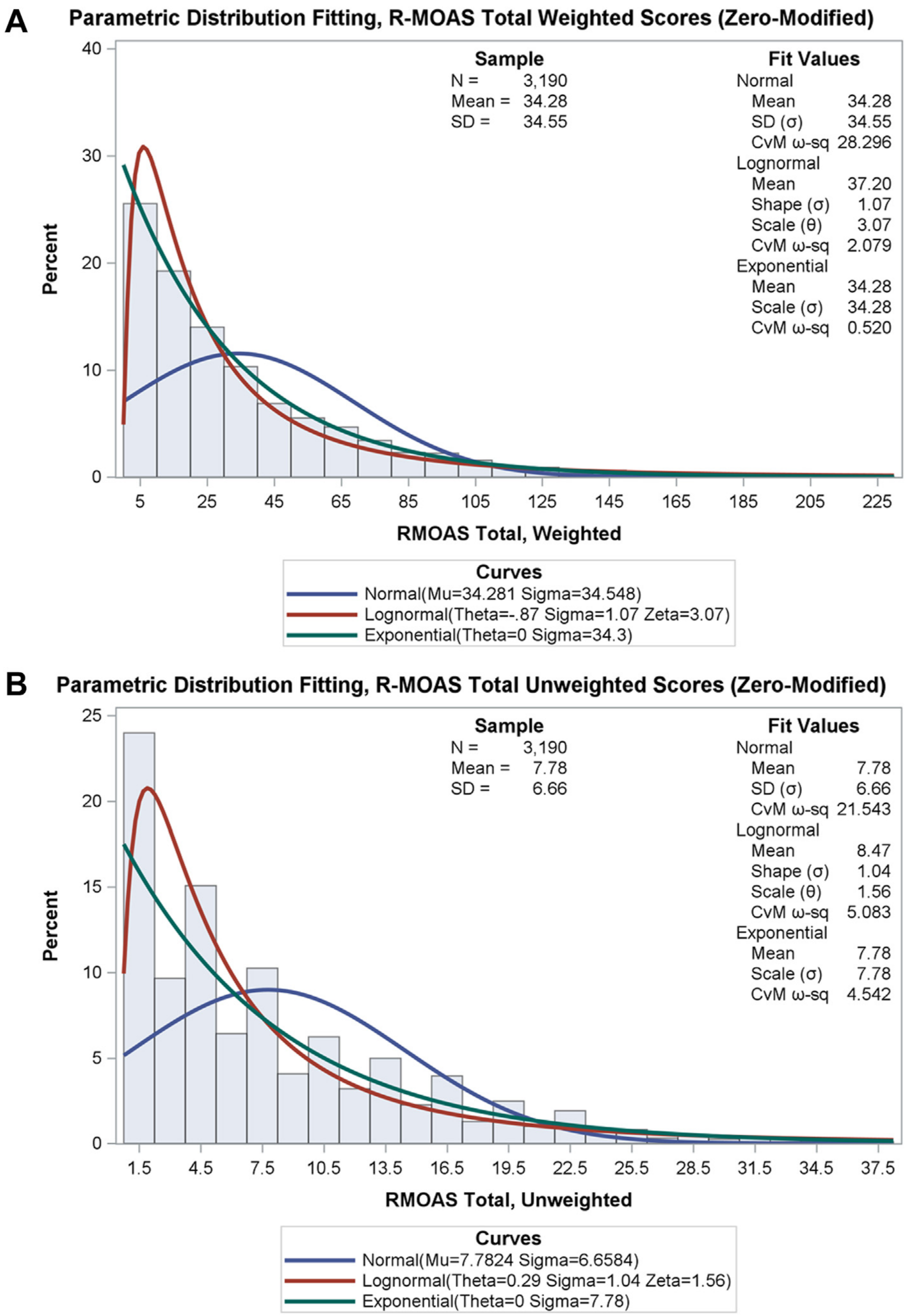
There were 686 patients who had an office visit within 30 days of their intake evaluation, at which R-MOAS and ConnGI-P data were obtained a second time. Total score test–retest reliability was 0.49 using weighted and 0.71 using unweighted scores. For comparison with an established rating scale, ρ for the Conners Global Index Total T score was 0.70. Bearing in mind that scores from the second assessment may have been affected by treatments that varied among patients, these estimates are likely lower-bound estimates of test–retest reliability of the measure itself (ie, what similarly spaced assessments with no external influences on the "true scores" would yield).

Item Response Theory Assessment

Table 2 contains IRT graded-response model parameter estimates for each item's response option thresholds and slopes. Items' option response functions (ORFs) are shown graphically in Figure S4, available online. The slope parameter represents the item's ability to distinguish between individuals who differ in the underlying trait (in this case, aggressive behavior, as measured by the R-MOAS). The range of slopes is 0.89 to 3.00; items for the self-aggression items group were lower albeit still in the moderate range,[37, p 34] reflecting in part their lower correlations with scale totals. Excluding them, the smallest slope is 1.83, which is in the "very high" range.[37]

A threshold is the level of the underlying trait at which the response choice thought to indicate higher severity does in fact become more likely than the less severe adjacent choice. Thus, there are k-1 thresholds, where k is the number of choices. The ORFs in Figure S4, available online, mostly show the expected form of ogives for the highest and lowest items, whereas the intermediate ones have unimodal curves. The exceptions are the highest self-harm items, the thresholds of which are at very high levels of overall aggression. The total test information

FIGURE 1 Parametric Distribution Fitting, R-MOAS Total (Zero-Modified)

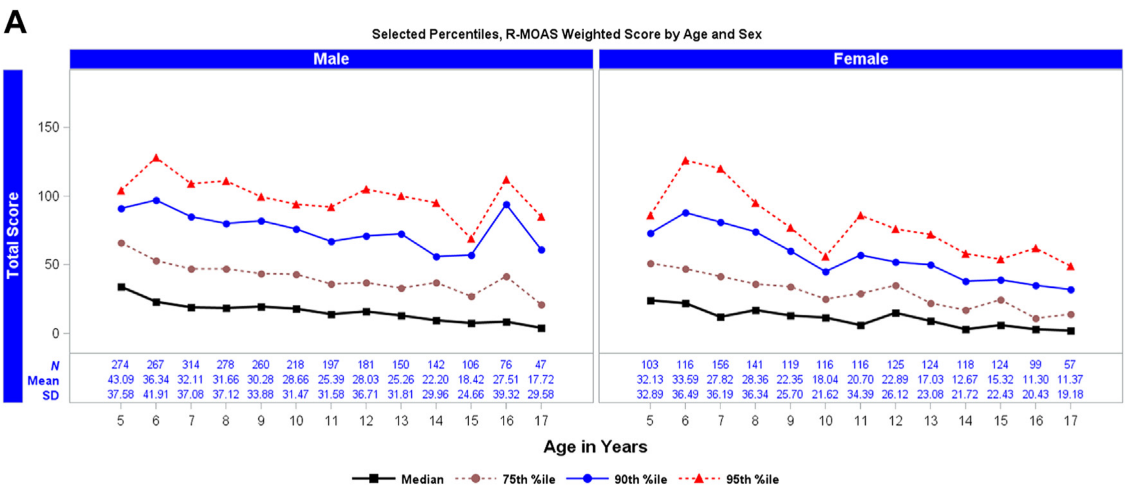


Note: Histogram of zero-modified R-MOAS Total scores (ie, excluding values of zero), weighted (b) and unweighted (b). Best-fit estimated curves and parameter values are shown for normal, lognormal, and exponential distributions. Distribution fit was assessed by the Cramér-von Mises ω^2 test, in which smaller values indicate better fit of observed scores with the respective parametric distribution density; the exponential distribution was the best match for the sample data. CvM = Cramér-von Mises; R-MOAS = Retrospective-Modified Overt Aggression Scale.

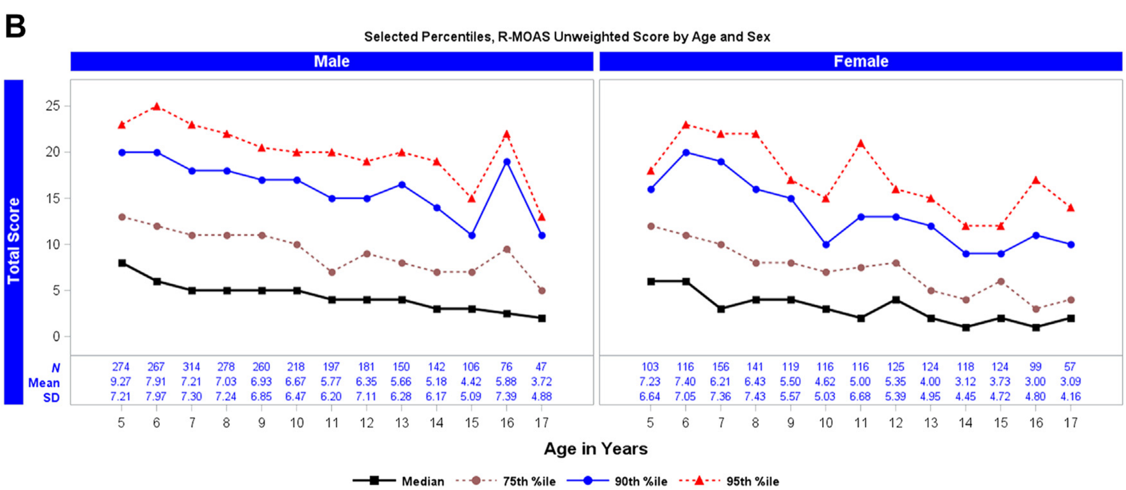
function (TIF) and item information functions (IIFs) are shown in Figure S5, available online. All but 3 self-harm items have information values over 1. Most curves' locations are shifted to the right, indicating relatively better performance differentiating among those with higher severity, a desirable property for use in clinical groups and for outcomes assessment. The lower item-total correlations for self-harm items that these curves reflect likely pertain to the factor structure, discussed in the next section.

Item response thresholds also enable comparisons between items to evaluate the severity with which the response is associated. Doing so permits assessment of the severity

FIGURE 2 Selected Percentiles, Retrospective-Modified Overt Aggression Scale (R-MOAS) Total Score by Sex and Age, Full Sample (N = 4,030)



Effect	B Estimate	df	*F*	*p*
Age (yrs)	-1.69	1, 4026	125.73	<.001
Sex				
Female	-5.27	1,4026	2.45	.12
Male (reference)				
Age x Sex	-0.14	1	0.20	.66



Effect	B Estimate	df	*F*	*p*
Age (yrs)	-0.35	1, 4026	139.96	<.001
Sex				
Female	-1.05	1,4026	2.48	.12
Male (reference)				
Age x Sex	-0.02	1,4026	.013	.72

Note: N = number of patients.

assumptions that the weighted scoring used in several OAS versions implies. Figure 3 shows threshold estimates for all 16 items. Within each item, there is the expected monotonic increase in threshold as response choices reflect higher severity. However, within aggressive behavior categories, the increases over items that the standard weighting

TABLE 1 Retrospective-Modified Overt Aggression Scale (R-MOAS) Reliability

		Weighted item scores				Unweighted frequency scores only			
	Sample size	Cronbach's α		McDonald's ω		Cronbach α's		McDonald's ω	
Internal consistency	**Full/ reduced**	**Full**	**Reduced**	**Full**	**Reduced**	**Full**	**Reduced**	**Full**	**Reduced**
Total score	4,030/3,453	0.84	0.83	0.87	0.85	0.88	0.87	0.90	0.89
Verbal	4,152/3,553	0.73	0.71	0.76	0.74	0.79	0.77	0.81	0.79
Physical	4,041/3,464	0.72	0.71	0.76	0.75	0.73	0.72	0.80	0.78
Property	4,044/3,467	0.76	0.74	0.77	0.75	0.77	0.76	0.82	0.80
Self-Directed	4,044/3,467	0.58	0.57	0.57	0.56	0.57	0.56	0.61	0.59
Temporal consistency	**Sample size**	**Baseline mean (SD)**	**Follow-up mean (SD)**		***ρ***	**Baseline mean (SD)**	**Follow-up mean (SD)**		***ρ***
Total score	686	27.02 (34.11)	20.80 (28.28)		0.69	6.11 (6.81)	5.01 (5.96)		0.71
Verbal	686	2.78 (4.25)	2.69 (4.21)		0.63	1.35 (1.86)	1.32 (1.84)		0.62
Physical	674	12.08 (17.68)	9.48 (14.66)		0.59	1.69 (2.24)	1.38 (1.91)		0.60
Property	675	8.01 (10.22)	5.63 (8.23)		0.64	2.30 (2.51)	1.68 (2.11)		0.63
Self-directed	665	4.49 (9.32)	3.29 (6.54)		0.49	0.85 (1.56)	0.70 (1.29)		0.50
Established comparator measure									
ConnGI, Total *T*	607					69.80 (15.80)	70.1 (16.00)		0.70

Note: *The "reduced" sample limited the proportion of R-MOAS Total scores of 0 to 10%. ConnGI = Conners Global Index; ρ = Spearman rank-order correlation coefficient.*
All values are significant with $p < .0001$.

approach assumes is not confirmed by the results. That is, the first 2 Physical items have identical rather than ascending thresholds.

Similarly, among Self-Directed items, Item SE1 ("…child pick[s] at or scratch[es] his or her skin, pull[s] out hair, or hit[s] himself or herself while upset or angry") has higher thresholds than Item SE2 ("child bang[s] his or her head, hit[s] his or her fists into the wall, or throw[s] himself or herself on the floor"), whereas the third (eg, self-cutting, burning) and fourth (severe injury, suicidal behavior) items' thresholds are nearly identical.

The implied severity hierarchy that increases through verbal → objects → self → physical aggression is not supported in the graded-response model. Notably, verbal aggression items have higher thresholds than their "level" counterparts in other categories. Self-Directed aggression items are, in fact, higher than the other-directed physical aggression ones.

Factor Structure

EFA scores from the discovery sample extracted 3 factors with eigenvalues greater than 1, corresponding to the "elbow" in the scree plot: 6.20, 1.58, and 1.03, which accounted for 39%, 10%, and 6% of variance, respectively. Comparison of 2- and 3-factor solutions showed a higher degree of cross-factor loadings with 3 factors, and the 2-factor solution was more readily interpretable. Parallel analysis with 1,000 iterations[36] confirmed the suitability of a 2-factor solution (Figure S6, available online). Table S2, available online, contains the rotated factors' standardized regression coefficients.

The first factor (Figure 4) comprises verbal lashing out and threats, plus the 2 least severe forms of physical and property-related behaviors. This constellation may be summarized as "Eruptive," with chiefly verbal and minor physical behaviors. The second factor involves more harmful behaviors directed toward oneself and property that culminate in injury or damage. For these items taken as a group, "Harmful/Distressed" may be an apt description of this factor. The rotated solution yields a correlation between factors of 0.40. The 2 most severe Physical items loaded equally on both factors, and so were not included in further assessment of factor structure.

In CFA, the factor pattern identified in EFA (Table 3) was applied to the confirmatory sample. It shows acceptable fit to the observed data based on standard metrics (Bentler comparative fit index [CFI] = 0.93, root mean square error of approximation [RMSEA] = 0.07). The 4-factor model based on item categories is similar (CFI = 0.93, RMSEA = 0.06), but with higher (worse) information criteria results

TABLE 2 Item Response Theory Properties

		Unidimensional		Multidimensional		
Item no.	**Item content**	**Parameter**	**Value (SE)**	**Parameter**	**Factor 1 Value (SE)**	**Factor 2 Value (SE)**
1	Verbal, brief outburst	Threshold 1	0.33 (0.03)			
		Threshold 2	1.53 (0.05)	Factor loading	0.76 (0.02)	
		Slope	1.86 (0.06)	Discrimination parameter	1.97 (0.07)	
2	Verbal, sustained outburst, <5 min	Threshold 1	0.68 (0.03)			
		Threshold 2	1.75 (0.05)	Factor loading	0.82 (0.02)	
		Slope	2.27 (0.08)	Discrimination parameter	2.44 (0.10)	
3	Verbal, sustained outburst, >5 min	Threshold 1	0.99 (0.03)			
		Threshold 2	1.85 (0.05)	Factor loading	0.82 (0.02)	
		Slope	2.37 (0.09)	Discrimination parameter	2.45 (0.10)	
4	Verbal, threats	Threshold 1	1.36 (0.04)			
		Threshold 2	2.38 (0.07)	Factor loading	0.78 (0.02)	
		Slope	2.15 (0.10)	Discrimination parameter	2.09 (0.10)	
5	Property, minor	Threshold 1	-0.36 (0.02)			
		Threshold 2	0.60 (0.02)			
		Threshold 3	1.31 (0.03)	Factor loading	0.85 (0.01)	
		Slope	2.70 (0.08)	Discrimination parameter	2.76 (0.09)	
6	Property, moderate	Threshold 1	0.01[‡](0.02)		0.00 (0.07)	
		Threshold 2	0.80 (0.03)		-2.28 (0.09)	
		Threshold 3	1.44 (0.03)	Factor loading	0.87 (0.01)	
		Slope	3.00 (0.10)	Discrimination parameter	3.04 (0.11)	
7	Property, destructive	Threshold 1	1.22 (0.04)			
		Threshold 2	1.93 (0.06)			
		Threshold 3	2.43 (0.08)	Factor loading		0.86 (0.02)
		Slope	2.15 (0.09)	Discrimination parameter		2.82 (0.17)
8	Property, fire, or weaponized object	Threshold 1	1.85 (0.06)			
		Threshold 2	2.56 (0.09)			
		Threshold 3	3.15 (0.13)	Factor loading		0.86 (0.02)
		Slope	2.24 (0.12)	Discrimination parameter		2.82 (0.20)
9	Self: pick, scratch, hit	Threshold 1	1.31 (0.05)			
		Threshold 2	2.28 (0.09)			
		Threshold 3	3.01 (0.12)	Factor loading		0.62 (0.03)
		Slope	1.23 (0.06)	Discrimination parameter		1.35 (0.07)
10	Self: bang head, punch wall[a]	Threshold 1	0.99 (0.03)			
		Threshold 2	1.70 (0.05)			
		Threshold 3	2.50 (0.08)			
		Slope	1.81 (0.07)			
11	Self: cut, bruise, burn	Threshold 1	3.63 (0.29)			
		Threshold 2	5.10 (0.44)			
		Threshold 3	6.09 (0.55)	Factor loading		0.57 (0.03)
		Slope	0.89 (0.08)	Discrimination parameter		1.19 (0.09)
12	Self: severe injury or suicide attempt	Threshold 1	3.85 (0.34)			
		Threshold 2	5.27 (0.52)			
		Threshold 3	6.25 (0.66)	Factor loading		0.61 (0.07)
		Slope	1.06 (0.12)	Discrimination parameter		1.33 (0.14)

(continued)

TABLE 2 Continued

		Unidimensional		Multidimensional		
Item no.	**Item content**	**Parameter**	**Value (SE)**	**Parameter**	**Factor 1 Value (SE)**	**Factor 2 Value (SE)**
13	Physical, gestured to strike	Threshold 1	0.23 (0.02)			
		Threshold 2	1.12 (0.03)			
		Threshold 3	1.85 (0.05)	Factor loading	0.81 (0.03)	
		Slope	2.34 (0.08)	Discrimination parameter	2.32 (0.08)	
14	Physical, no injury	Threshold 1	0.17 (0.02)			
		Threshold 2	1.04 (0.03)			
		Threshold 3	1.75 (0.05)	Factor loading	0.79 (0.03)	
		Slope	2.30 (0.07)	Discrimination parameter	2.23 (0.08)	
15	Physical, mild injury[a]	Threshold 1	1.23 (0.04)			
		Threshold 2	2.13 (0.06)			
		Threshold 3	2.72 (0.09)			
		Slope	2.25 (0.09)			
16	Physical, serious injury[a]	Threshold 1	2.51 (0.11)			
		Threshold 2	3.29 (0.17)			
		Threshold 3	3.99 (0.25)			
		Slope	1.83 (0.13)			

Note: *SE = standard error.*
[a]*Items that did not load differentially on either factor in exploratory factor analyses (EFA).*

(Akaike Information Criterion [AIC]: 643 vs 857; Schwarz Bayesian Information Criterion [SBIC]: 783 vs 1070). Both are superior to a single common-factor model (CFI = 0.89 and poorer information criteria performance). The correlation of the 2 factors in the CFA solution is 0.77, which is higher than in the EFA and probably due to fitting differences related to axis rotation.[38] The 2-factor model omits items that did not load preferentially on either factor in the CFA model, which may have improved fit statistics somewhat; however, testing this 12-item 2-factor model vs a 12-item single-factor model still shows vastly better fit for the former (χ^2, 593 vs 1166; SBIC 783 vs 1365; CFI = 0.93 vs 0.92). On balance, the 2-factor solution has the advantages of stronger information criteria metrics and parsimony relative to the 4-factor solution, as well as being preferred in the EFA assessment for the optimal number of factors.

Temporal stability of factors was evaluated via Spearman ρ using the same observations as in the overall test–retest reliability assessment. Reliability for factor 1 is 0.70 and for factor 2 is 0.52, mirroring the lower reliability for self-directed items that factor 2 emphasizes.

Item Response Multidimensional Theory Model

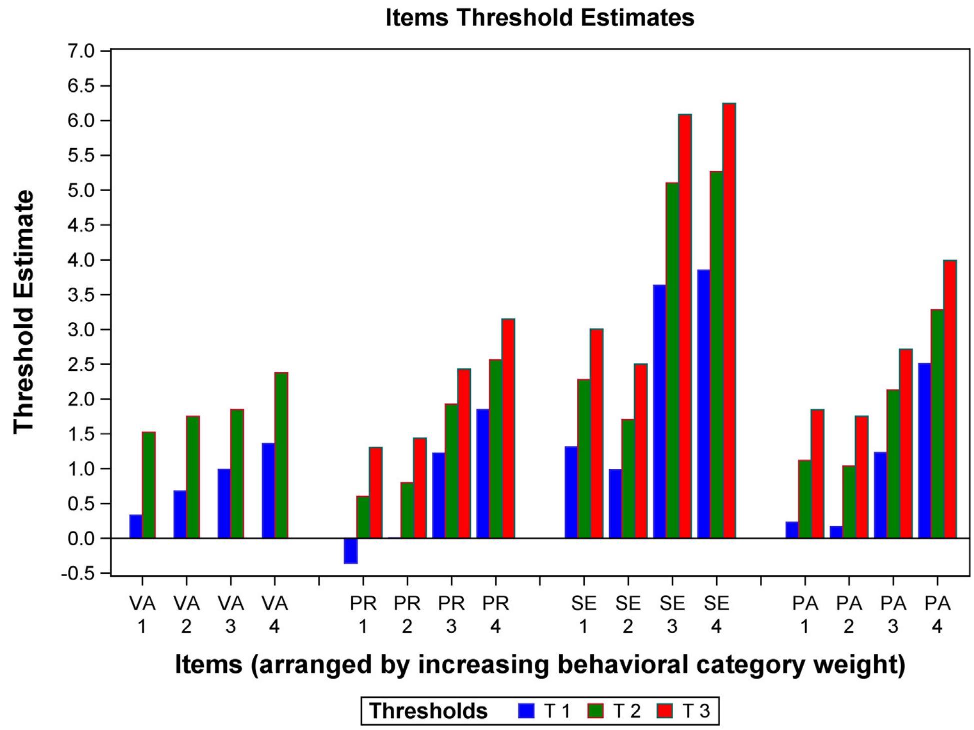
Because factor analyses indicate better fit for a two-factor solution over a single common-factor model, a multidimensional graded-response model IRT was evaluated. IRT produces a score (θ) similar to a standard score based on the underlying score distribution. The discrimination parameters (slopes) and factor loading of IRT-derived θ scores appear in the "Multidimensional" section on the right side of Table 2. As expected, fit is better relative to the unidimensional graded-response model: SBIC 63,642 vs 73,539; M_2 4,206 vs 6,570; RMSEA both 0.04. Correlations of IRT-derived θ scores with other measures used to assess validity are presented in Table S3, available online, and results confirm similarity with the correlations with factors using raw scores.

Validity

The correlations between R-MOAS scores and other established measures administered at intake were used to evaluate convergent and discriminant validity. Table 4 presents Spearman correlations for R-MOAS weighted and unweighted scores.

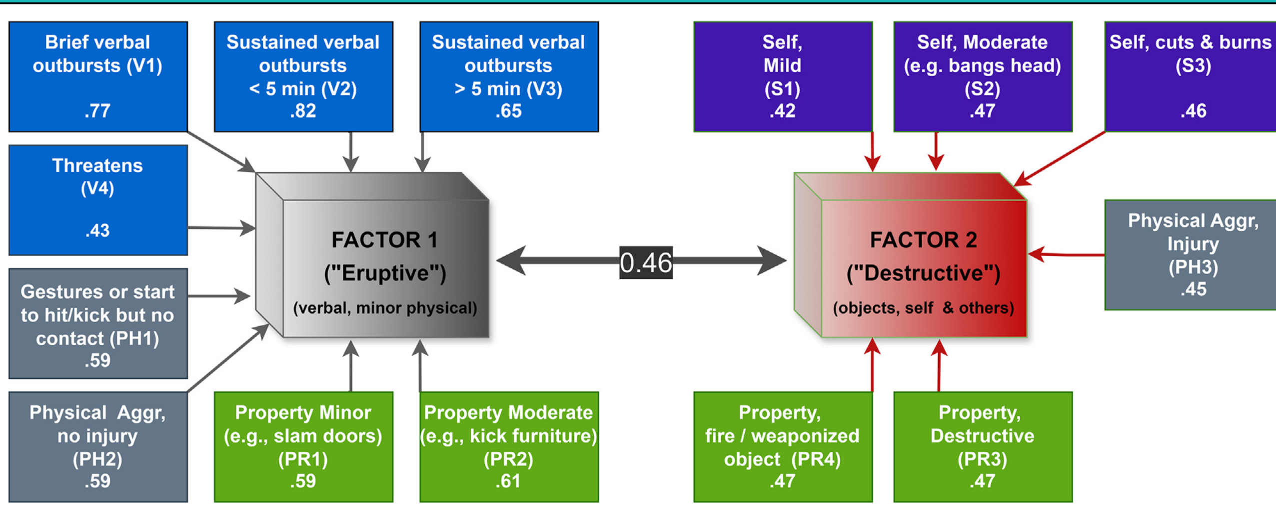
Moderate correlations ($r > 0.50$) were found between weighted R-MOAS Total score and measures of constructs expected to show positive associations, supporting convergent validity, notably those with ConnGI Total (0.61) and Emotional Lability (0.63) T scores, and the CBCL Externalizing scales (0.68). Correlations using unweighted R-MOAS scores were usually higher than those using weighted scores. Smaller associations with measures less strongly related to aggressive behavior supported discriminant validity, including CBCL Internalizing (0.35) and

FIGURE 3 Item Response Theory Threshold Estimated for Graded-Response Model



Note: *Threshold parameter estimates from Table 2 and Figure S3, available online, are arranged graphically to evaluate assumptions of the R-MOAS weighted scoring system. The threshold between one response choice and the next is the amount of the underlying trait (in this case, total aggressive behavior score) where the sample has an equal probability of endorsing either one. It is the point in the underlying trait continuum at which the next more "severe" response choice becomes more likely. The R-MOAS scoring approach gives larger weights (1) to later items in each behavioral category and (2) to the categories themselves (see text). Under the first assumption, each item's cluster of bars in this figure should increase to the next one. This does not occur consistently within the Property-Directed and Self-Directed categories. Under the second assumption, bars should increase across categories from left to right. It is undermined by the Property-Directed group of items being lower than Verbal, and the Physical items being lower than the Self-Directed category. R-MOAS = Retrospective-Modified Overt Aggression Scale.*

FIGURE 4 Exploratory Factor Analysis of the R-MOAS, Two-Factor Solution



Note: *Path diagram for exploratory factor analysis after oblique varimax rotation, showing factors with loadings >0.40 (n = 2,015).*

TABLE 3 Confirmatory Factor Analysis (N = 2,015)

Fit statistics	2 factors based on EFA	4 factors based on item categories	1 Common factor
RMSEA	0.07	0.06	0.06
CFI	0.93	0.93	0.89
χ^2 (df)	593 (53)	781 (98)	1,845 (104)
AIC	643	857	1,144
SBIC	783	1,070	2,111

Item loadings	Factor 1: "eruptive"	Factor 2: harmful/ distressed"	Factor 1: verbal	Factor 2: property	Factor 3: self	Factor 4: hysical	Single factor
Verbal_1_Brief	0.77		0.77				0.73
Verbal_2_Under5m	0.84		0.85				0.80
Verbal_3_Over5m	0.72		0.70				0.71
Verbal_4_Threat	0.59		0.56				0.60
Propty_1_Minor	0.85			0.85			0.84
Propty_2_Moder	0.89			0.90			0.90
Propty_3_Destruc		0.77		0.53			0.60
Propty_4_Severe		0.55		0.29			0.44
SlfHrm_1_Minor		0.50			0.53		0.46
SlfHarm2_Moder[a]					0.78		0.63
SlfHrm_3_Injur[b]					−0.08		0.10
SlfHrm_4_Severe		0.20			0.02		0.15
Physcl_1_Gesture	0.83					0.83	0.80
Physcl_2_NoInjur	0.78					0.84	0.81
Physcl_3_MinInjur[a]						0.60	0.62
Physcl_4_SigInjury[a]						0.13	0.24

Note: AIC = Akaike information criterion; CFI = Bentler comparative fit index; df = degrees of freedom; EFA = exploratory factor analysis; RMSEA = root mean square error of approximation; SBIC = Schwarz Bayesian information criterion.
[a]*Item from EFA with loading of zero in confirmatory factor analysis; excluded from model test and fit statistics.*
[b]*Items that did not load differentially on either factor in EFA.*

Somatic Problems (0.18). Similar patterns were observed for the Verbal, Physical, and Property-Related subscales, whereas the Self-Directed subscale had lower correlations with all ConnGI and CBCL scales.

Table 4 also shows the correlations of validity measures with the factors identified for the 2-factor solution derived from CFA (Table S3, available online, shows these correlations based on IRT-derived scores). Both factors show moderate correlations with scales relating to impulsivity, emotional lability (per ConnGI), and externalizing behavior indices, with factor 2 values consistently lower than those of factor 1. To shed further light on the latent variables that the 2 factors indicate, factor scores using the IRT-derived scores were examined in a multivariable model with CBCL subscales and the ConnGI Restless/Inattentive subscale (ConnGI Emotional Lability subscale was omitted because of collinearity with internalizing scales) as predictors. In these joint prediction models (Table 5), the factors vary in the influence that several predictors exert on them. Although the model for factor 1 shows no significant effects of the 2 affect-related CBCL subscales, one (Anxious/Depressed) is a significant predictor for factor 2 and the other (Withdrawn/Depressed) nearly so ($p = .06$). Thought Problems also has a stronger relationship with factor 2 than factor 1 (partial $\omega^2 = 0.02$ vs 0.002).[39] On the other hand, externalizing scales have much stronger associations with factor 1 than 2 (CBCL Aggressive, partial ω^2 0.13 vs 0.01; Rule-Breaking Behavior, 0.16 vs .005). ConnGI Restless/Inattentive scores make a substantial unique contribution to factor 1, while having no such effect on factor 2. (The CBCL Attention Problems scale's coefficients are negative for both factors in multivariable models and would be found even if ConnGI-P Restless/ Inattentive were omitted. Among this scale's 10 items, only

TABLE 4 R-MOAS Validity: Intercorrelations With *T* Scores of Other Measures

	ConnGI Total	ConnGI Restless/ Impulsive	ConnGI Emot Lability	CBCL Total	CBCL Externalizing	CBCL Internalizing	CBCL Anxious/ Depressed	CBCL Depressed/ Withdrawn	CBCL Somatic Probs	CBCL Social Probs	CBCL Thought Prob	CBCL Atten Prob	CBCL Aggr Behav	CBCL Rule Breaking
Descriptive/univariate														
Mean (SD)	71.28 (15.76)	69.86 (15.78)	65.89 (16.78)	64.93 (9.57)	62.24 (10.69)	58.90 (10.20)	62.09 (10.17)	63.59 (10.27)	61.50 (9.44)	62.73 (9.34)	64.75 (8.93)	67.63 (11.26)	64.71 (9.54)	61.40 (8.29)
Validity/Spearman ρ														
R-MOAS Total														
Weighted	**0.61**	**0.50**	**0.63**	**0.62**	**0.68**	0.35	0.33	*0.29*	*0.18*	0.47	0.49	0.36	**0.61**	**0.67**
Unweighted	**0.64**	**0.53**	**0.67**	**0.64**	**0.71**	0.36	0.35	*0.30*	*0.19*	0.48	**0.50**	0.37	**0.63**	**0.70**
R-MOAS Verbal														
Weighted	**0.56**	0.46	**0.62**	**0.56**	**0.63**	0.33	0.31	*0.28*	*0.18*	0.42	0.42	*0.28*	**0.55**	**0.63**
Unweighted	**0.57**	0.47	**0.62**	**0.55**	**0.63**	0.33	0.31	*0.28*	*0.18*	0.41	0.42	*0.28*	**0.54**	**0.63**
R-MOAS Physical														
Weighted	**0.50**	0.42	0.48	0.49	**0.57**	*0.24*	*0.22*	*0.21*	*0.09*	0.38	0.38	*0.30*	**0.53**	**0.55**
Unweighted	**0.51**	0.43	0.49	**0.50**	**0.58**	*0.24*	*0.22*	*0.21*	*0.09*	0.38	0.38	0.31	**0.53**	**0.51**
R-MOAS Property														
Weighted	**0.54**	0.45	**0.57**	**0.54**	**0.60**	*0.29*	*0.28*	*0.24*	*0.16*	0.40	0.40	0.32	**0.54**	**0.54**
Unweighted	**0.55**	0.45	**0.58**	**0.54**	**0.61**	*0.29*	*0.28*	*0.24*	*0.16*	0.41	0.40	0.32	**0.54**	**0.55**
R-MOAS Self														
Weighted	0.35	*0.28*	0.38	0.43	0.39	0.33	0.32	*0.27*	*0.20*	0.32	0.40	*0.24*	0.35	0.38
Unweighted	0.36	*0.29*	0.38	0.44	0.39	0.33	0.32	*0.27*	*0.20*	0.33	0.40	*0.25*	0.36	0.38
Factor 1[a] ("Eruptive")	**0.57**	**0.51**	**0.60**	**0.57**	**0.64**	0.32	0.29	*0.25*	*0.14*	0.48	0.45	0.34	**0.67**	**0.61**
Factor 2[a] ("Harmful/Distressed")	0.39	0.30	**0.34**	**0.39**	**0.39**	*0.25*	*0.24*	*0.21*	*0.15*	0.38	0.38	*0.26*	0.42	0.40
Intercorrelations of validity-measures/Spearman r														
ConnGI Total	1	0.94	0.78	0.75	0.75	0.41	0.38	0.29	0.21	0.57	0.61	0.66	0.66	0.73
ConnGI Restless/impulsive		1	0.59	0.68	0.67	0.32	0.31	0.23	0.18	0.52	0.58	0.69	0.6	0.65
ConnGI Emotional Lability			1	0.69	0.69	0.53	0.49	0.41	0.28	0.53	0.52	0.41	0.57	0.73
CBCL Total				1	0.84	0.75	0.68	0.59	0.5	0.77	0.78	0.67	0.74	0.81
CBCL Externalizing					1	0.45	0.39	0.35	0.24	0.61	0.6	0.54	0.91	0.91
CBCL Internalizing						1	0.87	0.81	0.52	0.6	0.55	0.37	0.34	0.48
CBCL Anxious/Depressed							1	0.60	0.54	0.62	0.55	0.36	0.29	0.46
CBCL Depressed/Withdrawn								1	0.48	0.47	0.43	0.30	0.27	0.39
CBCL Somatic Problems									1	0.38	0.40	0.22	0.17	0.26
CBCL Social Problems											0.63	0.54	0.56	0.61
CBCL Thought Problems											1	0.56	0.55	0.58
CBCL Attention Problems												1	0.48	0.54
CBCL Aggressive Behavior													1	0.77

Note: *n = 3,129. Correlations involving CBCL and ConnGI scores use* T *scores (standardized by age and sex) of these measures. In the Validity section, correlations ≥0.50 are shown in boldface type, and correlations ≤0.30 are shown in italic type. Aggr = aggressive; Behav = behavior; CBLC = Child Behavior Checklist; ConnGI = Conners Global Index–Parent Version; probs = problems; R-MOAS = Retrospective Modified Overt Aggression Scale.*
All values, p *< .0001.*
[a]*Scores using factor coefficients from confirmatory factor analysis.*

TABLE 5 Regression of IRT Θs From Two-Factor Solution on CBCL and ConnGI Subscales

	Factor 1				Factor 2			
	Parameter (β) (95% CI)		t	p	Parameter (β) (95% CI)		t	p
ConnGI Restless/Inattentive	0.009	(0.06 to 0.11)	8.52	<.00	0.0001	(−0.02 to 0.002)	0.15	.88
CBCL Anxious/Depressed	0.002	(−0.03 to 0.026)	0.41	.68	0.006	(0.003 to 0.01)	3.72	.0002
CBCL Withdrawn/Depressed	0.002	(−0.001 to 0.04)	1.35	.18	0.003	(0.0004 to 0.006)	2.29	.02
CBCL Somatic Problems	−0.0005	(−0.003 to 0.02)	−0.38	.70	0.001	(−0.002 to 0.003)	0.66	.51
CBCL Social Problems	0.02	(0.001 to 0.006)	1.34	.19	−0.001	(−0.004 to 0.03)	−0.49	.62
CBCL Thought Problems	0.005	(0.001 to 0.008)	2.74	.006	0.02	(0.01 to 0.02)	8.57	<.00
CBCL Attention Problems	−0.01	(−0.13 to −0.09)	−10.14	<.00	−0.003	(−0.006 to −0.0006)	−2.43	.02
CBCL Rule-Breaking Behavior	0.014	(0.01 to 0.18)	7.76	<.00	0.012	(0.01 to 0.015)	6.24	<.00
CBCL Aggressive Behavior	0.05	(0.046 to 0.055)	22.69	<.00	0.015	(0.01 to 0.02)	6.64	<.00
Sex (0 = male, 1 = female)	−0.10	(−0.14 to −0.05)	−4.59	<.00	−0.09	(−0.13 to 0.05)	−4.20	<.001
Age, y	−0.02	(−0.024 to −0.01)	−4.69	<.00	−0.01	(−0.016 to −0.002)	−2.39	.02
Model Fit (F_{df}, p)	$F_{11,3117}$ = 333.57, p < .0001				$F_{11,3117}$ = 111.09, p <.0001			
Model Effect Size	R^2 = 0.54 (95% CI = 0.52 to 0.56)				R^2 = 0.28 (95% CI = 0.25 to 0.30)			

Note: CBCL = Child Behavior Checklist; ConnGI = Conners Global Index, Parent version; df = degrees of freedom; IRT = Item Response Theory.

1 item bears on the behavioral impulsivity; others items appear more connected with "cognitive disengagement"[40] [daydreaming, confusion, poor concentration, etc]). Although male participants have higher scores than female participants on factor 1, there is no significant sex difference on factor 2. Scores on both factors decline with age, albeit more steeply for factor 1 (partial $\omega^2 = 0.008$ vs 0.002). Taken together, these results suggest that factor 2 signals a phenotype with possibly greater affective distress, less generalized impulsiveness, and for whom the "protective anti-aggressive" effects of female sex and older age are either absent or reduced.

DISCUSSION

Development of the R-MOAS was motivated by the need for a measure of child and adolescent aggressive behavior that quantifies common specific forms of aggression while separating them from other behavioral difficulties and higher-order constructs that, albeit correlated, are distinct. This psychometric assessment of the R-MOAS in a large clinical cohort of individuals 5 to 17 years of age indicates that it meets standard criteria for internal consistency and validity. Subscales also demonstrated good internal consistency, although that for the Self subscale is lower. Test–retest reliability, estimated from a subsample who had another visit within 30 days of the initial administration, was acceptable and on par with another well-established parent-reported behavioral measure. Clinical trials for children's aggressive behavior also indicated that this measure is sensitive to treatment-related changes.[26,27]

The skewed score distribution matches that often observed for behavioral disturbances, where zero and lower scores predominate and progressively taper with higher scores that indicate greater severity. Nevertheless, data in Table S1, available online, indicate the high frequency, in absolute terms, of aggressive behavior in this group referred within primary care settings for evaluation of behavioral concerns. For example, 24% were reported to have struck another person in the prior week once or twice, and another 12% were said to have done so 3 or 4 times.

Kay *et al.* introduced a weighting scoring system for the Overt Aggression Scale—Modified, which successor versions maintained. This report is the first to compare weighted with unweighted scores, and results do not demonstrate superiority on psychometric attributes of the weighted scoring approach. Instead, tests of internal consistency and temporal stability reliabilities showed overall advantage for the unweighted scores. Correlations with other measures used to examine validity also favored unweighted scores, in part reflecting the smaller error variance that the reliability tests showed. Attempting to amplify the effect on total scores of more severe aggressive behaviors, such as physical combativeness over verbal lashing out, has logical appeal. But when small rating differences exert a large impact on highly weighted summary scores these totals might become less stable.

Moreover, using IRT to empirically evaluate the presumed gradients in severity by items within and across behavioral categories demonstrates that these assumptions are invalid. These findings indicate there is no justification for continued use of the weighted scoring system, at least with youth populations.

EFA suggested 2 factors that differ in the harmfulness and object of aggressive behavior. CFA largely confirmed the pattern and was overall better than a behavioral categories-as-factors model, and both were superior to a single common-factor model. Factor 1 contains Verbal items as well as lower-severity Property-Related and Physical items, but no Self-Directed items. Factor 2 contains Self-Directed items and higher severity Property-Related items but none from Verbal. Using other measures as predictors in a multivariable model, factor 2 ("Harmful/Distressed") showed much stronger associations with affective symptoms and was more stable over age and equally distributed between male and female participants. In contrast, factor 1 ("Eruptive") showed greater associations with impulse control and other externalizing behavior problems; factor 1 scores also decreased with age, and were more likely to be endorsed for male than female participants. Factor 2's smaller item loadings are consistent with a factor that generates a "signal" but whose items may aggregate less often. Although, as noted above, the association between self-directed and other forms of harmful behavior has been reported previously,[41] they may have different risk factors and other determinants that may be important to parse more thoroughly. The ability of the R-MOAS to "detect" and provide scoring coefficients for factor 2 is a desirable feature in psychiatric contexts. A better understanding of youth with self-injurious behavior and dysregulated affect is a high priority.

Several points should be kept in mind as potential limitations of this assessment of the R-MOAS. Because its data did not come from an epidemiological sample, this study does not furnish population norms for the R-MOAS. Also, the clinical settings from which these data originate represent essentially a convenience sample. How study findings generalize to other clinical settings is an open question. However, analyses from this clinical cohort leveraged the efficiency of a larger number of higher scores in the distribution, relative to those expected in a general population sample, to evaluate reliability, validity, and factor structure. The measure's potential usefulness in measuring aggressive behavior in youth might make it a suitable candidate for which to undertake the more costly effort of obtaining population-based estimates of its scores' frequency distributions.

Not all aggressive behavior arises in the context of behavioral health symptoms or concerns. Because findings on distribution, reliability, validity, and factor structure used data from a clinical sample, they may not generalize to youth who demonstrate similar behaviors but who are not identified in health care settings.

Item content was derived from a rating scale intended for adult psychiatric inpatients, for whom there are no published reports of more advanced psychometric evaluation (such as IRT and factor analysis). Given the dearth of instruments for youth focused on aggressive behavior per se, it was adapted for clinical and research use because it was recognized as an established measure and because starting *de novo* to develop and validate a new tool was infeasible. Despite the strengths of the R-MOAS seen in these analyses, alternative content developed specifically for children and adolescents might have even more favorable psychometric features.

Clinical Significance

The prominence of aggressive behavior, as a common and impairing target symptom in child and adolescent psychiatry necessitates a measurement tool that quantifies the severity of its specific manifestations. The R-MOAS offers a means to track this important outcome using a straightforward, adult-informant measure. In clinical use, it is recommended for consideration as part of routine assessment where there is a likelihood that aggressive behavior may be a presenting concern.

The behavioral categories of the R-MOAS are clinically useful for conveying the specific types of aggressive behaviors. It is worthwhile to know that a youth is predominantly verbally hostile and throws things, for example, while not being physically assaultive toward others. These subscales may not represent "factors," at least in the sense of indicating latent traits. Given the obvious similarities in their contents, their clustering (per CFA) might better represent "principal components." By comparison, the 2-factor solution shows clustering of items across behavioral categories that may indicate latent traits – impulsive, eruptive, temper loss (factor 1), and distress driven by primary affective disturbances with less influence of generalized impulse control issues (factor 2).

Although aggression is a transdiagnostic construct, treatment approaches remain strongly influenced by the overall psychiatric context in which it arises for individual patients. Concurrent assessment of the developmental, symptomatic, and environmental features that create susceptibility to aggressive behavior is therefore vital. Ideally, R-MOAS administration would be accompanied by administration of other tools that gauge the common contributors to aggression, including impaired impulse control (as in attention-deficit/hyperactivity disorder), emotional overreactivity (such as irritability), other mood and anxiety-related disturbances, other forms of disruptive behavior, developmental status, traumatic exposures, and features of the patient's interpersonal and material environments.[42] In many contexts, it is also valuable to include systematic assessments of motivational aspects of aggressive behavior useful for treatment planning. These include the following: (1) the degree to which patients' behaviors are distributed along the proactive/volitional–reactive/impulsive dimension,[43] and (2) overall pro-sociality vs callous disregard for others.[44]

Measurement-based care is desirable for most psychiatric conditions to ensure the effectiveness, timeliness, and tolerability of treatments. The R-MOAS is useful for this outcome monitoring role when aggressive behaviors are the focus of clinical attention. Youth treated for volatile and aggressive behavior often receive pharmacotherapy involving several agents, and rigorous measurement of treatment response is essential to optimizing their value while minimizing the adverse effects that can accompany these regimens.

CRediT authorship contribution statement

Joseph C. Blader: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

This article is part of a special series devoted to addressing aggressive behavior as a focus of psychiatric attention and how its manifestations and treatment needs may vary across psychiatric disorders. The series is edited by Guest Editor Joseph Blader, PhD, Deputy Editor Robert Findling, MD, MBA, and Editor Manpreet K. Singh, MD, MS.

Accepted March 6, 2024.

This article was reviewed under and accepted by Ad Hoc Editor Benedetto Vitiello, MD.

Joseph C. Blader is with the University of Texas Health Science Center at San Antonio.

The author has reported funding for this work from the Texas Health and Human Services Commission (085144601.2.2 to Dr. Blader).

Joseph C. Blader served as the statistical expert for this research.

The contributions of PROXIMA clinicians, staff, and participating families in acquiring the data for this report are gratefully acknowledged.

Disclosure: Joseph C. Blader has received research support from the Texas Child Mental Health Care Consortium and consultant's honoraria from Supernus Pharmaceuticals.

Correspondence to Joseph C. Blader, PhD, Department of Psychiatry and Behavioral Sciences, University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, MS 7719, San Antonio, TX 78229; e-mail: blader@uthscsa.edu

2949-7329/© 2024 The Author. Published by Elsevier Inc. on behalf of American Academy of Child & Adolescent Psychiatry. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

https://doi.org/10.1016/j.jaacop.2024.02.004

REFERENCES

1. Hoffmann JA, Stack AM, Monuteaux MC, Levin R, Lee LK. Factors associated with boarding and length of stay for pediatric mental health emergency visits. Am J Emerg Med. 2019;37(10):1829-1835. https://doi.org/10.1016/j.ajem.2018.12.041
2. Walker S, Barnett P, Srinivasan R, Abrol E, Johnson S. Clinical and social factors associated with involuntary psychiatric hospitalisation in children and adolescents: a systematic review, meta-analysis, and narrative synthesis. Lancet Child Adolesc Health. 2021;5(7):501-512. https://doi.org/10.1016/S2352-4642(21)00089-4
3. Tossone K, Jefferis E, Bhatta MP, Bilge-Johnson S, Seifert P. Risk factors for rehospitalization and inpatient care among pediatric psychiatric intake response center patients. Child Adolesc Psychiatry Ment Health. 2014;8(1):27. https://doi.org/10.1186/1753-2000-8-27
4. Kreider AR, Matone M, Bellonci C, *et al*. Growth in the concurrent use af antipsychotics with other psychotropic medications in Medicaid-enrolled children. J Am Acad Child Adolesc Psychiatry. 2014;53(9):960-970. https://doi.org/10.1016/j.jaac.2014.05.010
5. Loy JH, Merry SN, Hetrick SE, Stasiak K. Atypical antipsychotics for disruptive behaviour disorders in children and youths. Cochrane Database Syst Rev. 2017;8:CD008559. https://doi.org/10.1002/14651858.CD008559.pub3
6. Toteja N, Gallego JA, Saito E, *et al*. Prevalence and correlates of antipsychotic polypharmacy in children and adolescents receiving antipsychotic treatment. Int J Neuropsychopharmacol. 2014;17(7):1095-1105. https://doi.org/10.1017/S1461145712001320
7. Tremblay RE. Why socialization fails: the case of chronic physical aggression. In: Lahey BB, Moffitt TE, Caspi A, eds. Causes of Conduct Disorder and Juvenile Delinquency. Guilford Press; 2003:182-224.
8. Saylor KE, Amann BH. Impulsive aggression as a comorbidity of attention-deficit/hyperactivity disorder in children and adolescents. J Child Adolesc Psychopharmacol. 2016;26(1):19-25. https://doi.org/10.1089/cap.2015.0126
9. Zimmer-Gembeck MJ, Nesdale D, Webb HJ, Khatibi M, Downey G. A longitudinal rejection sensitivity model of depression and aggression: unique roles of anxiety, anger, blame, withdrawal and retribution. J Abnorm Child Psychol. 2016;44(7):1291-1307. https://doi.org/10.1007/s10802-016-0127-y
10. Freitag GF, Grassie HL, Jeong A, *et al*. Systematic review: questionnaire-based measurement of emotion dysregulation in children and adolescents. J Am Acad Child Adolesc Psychiatry. 2023;62(7):728-763. https://doi.org/10.1016/j.jaac.2022.07.866
11. Krahé B. The Social Psychology of Aggression. 3rd ed. Routledge; 2021.
12. Suris A, Lind L, Emmett G, Borman PD, Kashner M, Barratt ES. Measures of aggressive behavior: overview of clinical and research instruments. Aggress Violent Behav. 2004;9(2):165-227. https://doi.org/10.1016/S1359-1789(03)00012-0
13. O'Donnell O, House A, Waterman M. The co-occurrence of aggression and self-harm: systematic literature review. J Affect Disord. 2015;175:325-350. https://doi.org/10.1016/j.jad.2014.12.051
14. Barzman DH, Brackenbury L, Sonnier L, *et al*. Brief Rating of Aggression by Children and Adolescents (BRACHA): development of a tool for assessing risk of inpatients' aggressive behavior. J Am Acad Psychiatry Law. 2011;39(2):170-179; https://jaapl.org/content/39/2/170
15. Coccaro EF. The Overt Aggression Scale Modified (OAS-M) for clinical trials targeting impulsive aggression and intermittent explosive disorder: validity, reliability, and correlates. J Psychiatr Res. 2020;124:50-57. https://doi.org/10.1016/j.jpsychires.2020.01.007
16. Yudofsky SC, Silver JM, Jackson W, Endicott J, Williams D. The Overt Aggression Scale for the objective rating of verbal and physical aggression. Am J Psychiatry. 1986;143(1):35-39. https://doi.org/10.1176/ajp.143.1.35
17. Achenbach TM, Rescorla LA. Manual for the ASEBA School-Age Forms & Profiles. Youth, and Families: University of Vermont Research Center for Children; 2001.
18. Aman M, Leone S, Lecavalier L, Park L, Buican B, Coury D. The Nisonger Child Behavior Rating Form: Typical IQ Version. Int Clin Psychopharmacol. 2008;23(4):232-242. https://doi.org/10.1097/YIC.0b013e3282f94ad0
19. Reynolds CR, Kamphaus RW. Behavior Assessment System for Children. 3rd ed. Pearson Assessments; 2015.
20. Buss AH, Warren WL. Aggression Questionnaire (AQ) Manual. Western Psychological Services; 2000.
21. Langfus JA, Youngstrom EA, DuBois CM, Findling RL, Stepanova E. Scoping review: evidence-based assessment of reactive aggression in children. JAACAP Open. 2023;1(4):246-262. https://doi.org/10.1016/j.jaacop.2023.08.005

22. Vitiello B, Behar D, Hunt J, Stoff D, Ricciuti A. Subtyping aggression in children and adolescents. J Neuropsychiatry Clin Neurosci. 1990;2(2):189-192. https://doi.org/10.1176/jnp.2.2.189
23. Raine A, Dodge K, Loeber R, *et al*. The Reactive–Proactive Aggression Questionnaire: differential correlates of reactive and proactive aggression in adolescent boys. Aggress Behav. 2006;32(2):159-171. https://doi.org/10.1002/ab.20115
24. Halperin JM, McKay KE, Newcorn JH. Development, reliability, and validity of the Children's Aggression Scale—Parent Version. J Am Acad Child Adolesc Psychiatry. 2002;41:245-252. https://doi.org/10.1097/00004583-200203000-00003
25. Kay SR, Wolkenfeld F, Murrill LM. Profiles of aggression among psychiatric patients. I. Nature and prevalence. J Nerv Ment Dis. 1988;176(9):539-546. https://doi.org/10.1097/00005053-198809000-00007
26. Blader JC, Schooler NR, Jensen PS, Pliszka SR, Kafantaris V. Adjunctive divalproex versus placebo for children with ADHD and aggression refractory to stimulant monotherapy. Am J Psychiatry. 2009;166(12):1392-1401. https://doi.org/10.1176/appi.ajp.2009.09020233
27. Blader JC, Pliszka SR, Kafantaris V, *et al*. Stepped treatment for attention-deficit/hyperactivity disorder and aggressive behavior: a randomized, controlled trial of adjunctive risperidone, divalproex sodium, or placebo after stimulant medication optimization. J Am Acad Child Adolesc Psychiatry. 2021;60(2):236-251. https://doi.org/10.1016/j.jaac.2019.12.009
28. Conners CK. Conners 3rd Edition Manual. Multi-Health Systems. 2008.
29. SAS Institute Inc. SAS 9.4 User's Guide. 2nd ed. SAS Institute; 2016.
30. Cai L. flexMIRT: Flexible Multilevel Multidimensional Item Analysis and Test Scoring (Version 3.65) [computer program]. Vector Psychometric Group; 2022.
31. Dunn TJ, Baguley T, Brunsden V. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. Br J Psychol. 2014;105(3):399-412. https://doi.org/10.1111/bjop.12046
32. Hayes AF. OMEGA for SAS (Version 1.0) [Computer program]. Available at:. Accessed April 2, 2024. https://afhayes.com/spss-sas-and-r-macros-and-code.html
33. Samejima F. Estimation of latent ability using a response pattern of graded scores. Psychometrika. 1969;34(Suppl):1-97. https://doi.org/10.1007/BF03372160
34. de Ayala RJ. The influence of multidimensionality on the graded response model. Appl Psychol Meas. 1994;18(2):155-170. https://doi.org/10.1177/014662169401800205
35. Horn JL. A rationale and test for the number of factors in factor analysis. Psychometrika. 1965;30(2):179-185. https://doi.org/10.1007/BF02289447
36. Kabacoff RI. Determining the Dimensionality of Data: A SAS® Macro for Parallel Analysis. Available at:. Accessed April 2, 2024. https://support.sas.com/resources/papers/proceedings/proceedings/sugi28/090-28.pdf
37. Baker FB. The Basics of Item Response Theory. 2nd ed. ERIC Clearinghouse on Assessment and Evaluation; 2001. http://ericae.net/irt/baker
38. Marsh HW, Morin AJS, Parker PD, Kaur G. Exploratory structural equation modeling: an integration of the best features of exploratory and confirmatory factor analysis. Annu Rev Clin Psychol. 2014;10(1):85-110. https://doi.org/10.1146/annurev-clinpsy-032813-153700
39. Kroes ADA, Finley JR. Demystifying omega squared: practical guidance for effect size in common analysis of variance designs [published online ahead of print July 20, 2023]. Psychol Methods. https://doi.org/10.1037/met0000581
40. Becker SP. Systematic review: assessment of sluggish cognitive tempo over the past decade. J Am Acad Child Adolesc Psychiatry. 2021;60(6):690-709. https://doi.org/10.1016/j.jaac.2020.10.016
41. Barzman D, Hemphill R, Appel K, *et al*. A large naturalistic study on the BRACHA: vonfirmation of the predictive validity. Psychiatr Q. 2022;93(3):803-811. https://doi.org/10.1007/s11126-022-09993-4
42. Blader JC, Connor DF. Aggression in children: an integrative approach. In: Martin A, Volkmar FR, Bloch MH, eds. Lewis's Child and Adolescent Psychiatry: a Comprehensive Textbook. 5th ed. Wolters Kluwer; 2017:399-410.
43. Connor DF, Steingard RJ, Cunningham JA, Anderson JJ, Melloni RH Jr. Proactive and reactive aggression in referred children and adolescents. Am J Orthopsychiatry. 2004;74(2):129-136. https://doi.org/10.1037/0002-9432.74.2.129
44. Fontaine NM, McCrory EJ, Boivin M, Moffitt TE, Viding E. Predictors and outcomes of joint trajectories of callous-unemotional traits and conduct problems in childhood. J Abnorm Psychol. 2011;120(3):730-742. https://doi.org/10.1037/a0022620