

A Measure of Emotional Regulation and Irritability in Children and Adolescents: The Clinical Evaluation of Emotional Regulation–9

Journal of Attention Disorders
2020, Vol. 24(14) 2002–2011
© The Author(s) 2017



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1087054717737162
journals.sagepub.com/home/jad



Jenna Pylypow¹, Declan Quinn¹, Don Duncan², and Lloyd Balbuena¹

Abstract

Objective: To develop a scale for emotional regulation using item response theory. **Method:** Eighteen Swanson Nolan and Pelham (SNAP-IV) items that loaded on an emotional dysregulation factor were submitted to Rasch analysis. After eliminating the items that violated Rasch criteria, the remaining items were examined for reliability and validated against the Conners' emotional lability index. **Results:** A nine-item scale for emotional regulation was developed that satisfies the Rasch model and reliably distinguishes emotionally dysregulated/irritable children and adolescents. A score of 4 or higher in this scale has optimal accuracy for identifying children and adolescents with current significant dysfunction in emotional regulation. Among youth with ADHD inattentive, hyperactive–impulsive, and combined types, 42%, 56%, and 71% met the Clinical Evaluation of Emotional Regulation–9 (CEER-9) threshold for emotional lability, respectively. **Conclusion:** A nine-item scale whose sum total is a measure of emotional regulation is proposed as a tool for clinical and research purposes. (*J. of Att. Dis.* 2020; 24(14) 2002–2011)

Keywords

emotional regulation, emotional lability, irritability, item response theory, ADHD-associated problems, psychometrics

Introduction

Deficits in the self-regulation of emotions, or emotional dysregulation (EDr), have recently become a major focus of academic interest, but have been long identified and studied within the literature as many closely related or overlapping concepts, including negative emotionality, neuroticism, choleric or fiery temperament, and emotional or mood lability (Cole, Martin, & Dennis, 2004; Miller & Pilkonis, 2006; Widiger, 1998). Emotional dysregulation is the propensity for excessively and rapidly shifting emotions that is inappropriate to the situational context, age, and developmental stage (Cole, Michel, & Teti, 1994; Conklin, Bradley, & Westen, 2006; Glenn & Klonsky, 2009). It also encompasses irritability, which is described as the excessive physiological and negative affective reactivity to stimuli (Caprara et al., 1985; Rich et al., 2007) and is characterized by subjective feelings of anger, short/bad temper, crankiness, resentment, or annoyance—sometimes resulting in aggressive behavior (Russell A. Barkley & Benton, 1998; Barry, Marcus, Barry, & Coccaro, 2013; Caprara et al., 1985; Stringaris, 2011). Irritable children are more likely to drop out of school, be unemployed, and have strained adult relationships (Caspi, Wright, Moffitt, & Silva, 1998; Fergusson, John Horwood,

& Ridder, 2005; Kokko, Bergman, & Pulkkinen, 2003; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007; Sahl, Cohen, & Dasch, 2009). Emotionally impulsive children are also more likely to experience legal and financial problems in adulthood (R. A. Barkley & Fischer, 2010). Two prospective follow-up studies showed that adolescent irritability predicts future mood and anxiety disorders and suicidal behavior (Pickles et al., 2010; Stringaris, Cohen, Pine, & Leibenluft, 2009).

Poor emotional control is common, prevalent across the life span, and linked with many psychiatric and medical disorders. Although EDr is present in 3% to 20% of children and youth in general (Brotman et al., 2006), up to one third of psychiatric samples have irritability (Stringaris, 2011). Rates of EDr in ADHD vary from 8% to 80%, depending on the presence of comorbid disorders (Galanter et al., 2003;

¹University of Saskatchewan, Saskatoon, Canada

²BC Interior ADHD Clinic, Kelowna, British Columbia, Canada

Corresponding Author:

Declan Quinn, Child and Adolescent Psychiatry, Department of Psychiatry, University of Saskatchewan, Royal University Hospital, Saskatoon, Saskatchewan, Canada S7N 0W8.

Email: declan.quinn@usask.ca

Geller et al., 2001; Mick, Spencer, Wozniak, & Biederman, 2005). EDr is probably a core feature of oppositional defiant disorder (ODD; R. A. Barkley & Fischer, 2010; Martel, Gremillion, Roberts, von Eye, & Nigg, 2010; Pelham, Gnagy, Greenslade, & Milich, 1992; van Stralen, 2016). This is supported by the finding that emotionality as a toddler predicted ODD in childhood (Stringaris, Maughan, & Goodman, 2010). Children with ODD who display anger-irritability symptoms also tend to have more severe symptoms of anxiety and depression than ODD counterparts with behavioral symptoms (Drabick & Gadow, 2012). This implies that the affective symptoms of ODD may be of greater prognostic value for future mental disorders. EDr is also associated with many other childhood psychiatric disorders including bipolar disorder, depression, dysthymia, separation anxiety, and conduct disorder (*Diagnostic and Statistical Manual of Mental Disorders* [5th ed.; *DSM-5*; American Psychiatric Association [APA], 2013]; Birmaher, 2016; D. S. Shaw, Owens, Giovannelli, & Winslow, 2001). EDr is also associated with many adult psychiatric disorders including generalized anxiety disorder, premenstrual dysphoric disorder, posttraumatic stress disorder, substance-related and addictive disorders, and borderline personality disorder (APA, 2013; Winsper, Hall, Strauss, & Wolke, 2017). EDr is also common in many neuropsychiatric conditions including Huntington's disease (Demark & Gemeinhardt, 2002; Rosenblatt & Leroi, 2000), acquired brain injury (Demark & Gemeinhardt, 2002; Kim, Moles, & Hawley, 2001), and dementias (Cipriani, Vedovello, Nuti, & Di Fiorino, 2011; Kim et al., 2001). As such, it contributes to the overlap in diagnostic criteria and lack of specificity between psychiatric disorders (Caron & Rutter, 1991). For instance, disruptive mood dysregulation disorder (DMDD) was created, in part, because of the implausibly high rates of nonepisodic irritable children and youth misdiagnosed with bipolar disorder (Moreno et al., 2007; Roy, Lopes, & Klein, 2014). However, ODD and DMDD cannot be differentiated based on symptoms, thereby complicating the diagnostic process (Meyers, DeSerisy, & Roy, 2017). Clearly, the effective measurement of EDr in psychiatric disorders as well as medical disorders is important.

Distinguishing normal and pathological expressions of emotion depends critically on having valid and reliable instruments. The conceptual understanding and management of EDr can only be advanced with good measurement (Stringaris, 2011; van Stralen, 2016). Unfortunately, too many mental health rating scales take the validity of sum scores for granted (da Rocha, Chachamovich, de Almeida Fleck, & Tennant, 2013). Raw scores in many instruments are ordinal-level measures whose adjacent categories do not represent equal intervals (Hobart, Cano, Zajicek, & Thompson, 2007). Furthermore, the constructs underlying candidate items are often not well developed theoretically prior to scale development (Hobart et al., 2007). Within the

paradigm of classical test theory, the focus is on total scores (or average scores) in a scale, with the assumption that each question is weighted equally (Fayers & Machin, 2016). In effect, each item in the test is considered a "parallel test" and is assumed to be endorsed with the same frequency as the others (van Schuur, 2003). This is unlikely to hold in clinical settings. For example, among patients with neck pain, *having headaches* was more frequently endorsed than *being able to work* (van der Velde, Beaton, Hogg-Johnston, Hurwitz, & Tennant, 2009). Thus, a score of 10 might not reflect the same level of disability in people endorsing different sets of items. Newer psychometric techniques, such as Rasch modeling, a form of item response theory (IRT), can help answer the question whether the use of sum scores for a given test is justified or not. Sum scores are valid measures only when questionnaire responses conform to a probabilistic Guttman (staircase-like) pattern (Andrich, 1988). Otherwise, sum scores are misleading.

Many scales have been developed for the assessment of emotional regulation and related constructs. None has yet been developed for children and youth with the benefit of IRT, with a single exception. This is the Patient-Reported Outcomes Measurement Information System (PROMIS) Pediatric Anger Scale (Irwin et al., 2012). With six questions rated on a five-point scale, the instrument translates a participant's sum score into *T* scores (Irwin et al., 2012). The Difficulties in Emotion Regulation Scale (DERS) is based on a well-developed theory of emotion regulation that is contrasted with the constriction or control of emotions, leading to a six-factor instrument with satisfactory reliability and validity (Gratz & Roemer, 2004; Weinberg & Klonsky, 2009). Stringaris and colleagues designed a parent- and self-report scale to assess childhood and adolescent irritability, the affective reactivity index (ARI; Stringaris et al., 2012). While these scales each have their niche, a scale that is based on affective and behavioral traits such as those seen in ODD seems warranted. In this study, our objective was to develop a reliable and valid instrument to measure emotional regulation in children and youth using IRT—the current standard of test construction.

Method

Data and Measures

Our data came from a web-administered SNAP-IV 90-item rating scale site (www.adhdratingscales.com) managed by one of the authors (DD; Swanson, unpublished). Parents and teachers rated these scales after being referred to the site by 64 clinicians, 75% of whom were pediatricians. SNAP data from 3,374 children and youth were available for analysis. Previously, we factor analyzed the SNAP and found nine factors (Cavanagh, Quinn, Duncan, Graham, & Balbuena, 2014). The present study started with the factor

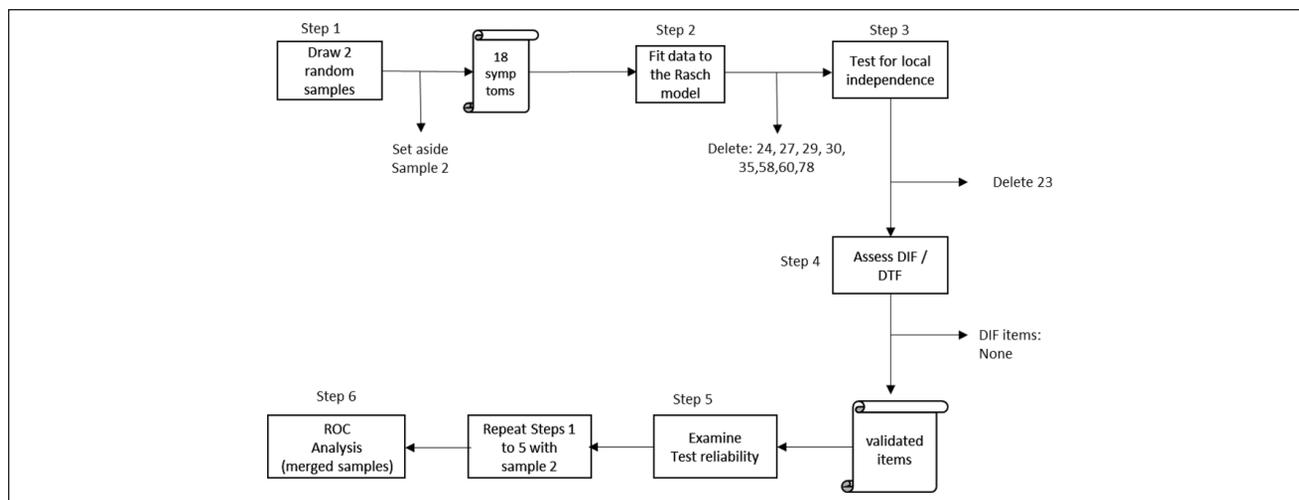


Figure 1. Schematic of analysis steps.

representing emotional dysregulation. The 18 SNAP items loading on emotional dysregulation were *loses temper; argues with adults; actively defies or refuses adult requests or rules; does things that annoy others; blames others for mistakes or misbehavior; touchy or easily annoyed; angry and resentful; spiteful or vindictive; quarrelsome; negative, defiant, disobedient, or hostile toward authority figures; uncooperative; acts smart; changes mood quickly and drastically; irritable; has excessive emotionality and attention-seeking behavior; unstable relationships with others; reactive mood; and impulsive*. These items were originally rated on a 0 to 3 Likert-type scale anchored on the extremes: *not at all* and *very much*. The item responses were dichotomized, with 0/1 recoded as 0, and 2/3 recoded as 1, to be amenable to Rasch analysis.

Overview of Rasch Analysis

In contrast to classical psychometric techniques, Rasch analysis calibrates a scale by matching item difficulties with test-taker abilities. By analogy to educational testing, a difficult question is one that is correctly answered by few test-takers compared with an easy question, which is correctly answered by many. Conversely, a test-taker who answers more items correctly is of higher ability than another who gets fewer correct items. Without loss of generality, one can interpret “higher ability” in the context of personality as somebody who has “more” of the trait in question. This way, both item “difficulties” and respondent “abilities” (trait levels) are calibrated with respect to each other. More formally, Rasch analysis models a test-taker’s probability of endorsing a symptom as a logistic function of the difference in that person’s trait level and the item’s difficulty (Fayers & Machin, 2016). If the pattern of test-taker responses is then found to conform, within tolerance, to a staircase-like

pattern (called a Guttman scale), then the count of symptoms in the scale suffices as a measure of the latent trait. In practice, this means that two children, each having a count of three symptoms, have the same level of emotional dysregulation regardless, of the particular symptoms endorsed.

Analytic Strategy

Our Rasch analysis consisted of six steps as depicted in Figure 1. First, we drew two random samples of 360 people who were matched for age, gender, reporter (parent or teacher) type, and raw score distribution. This was done because, in an overly large sample, small deviations can lead to an incorrect rejection of the hypothesis that the items fit a Guttman pattern—a Type 1 error (Martin-Lof, 1974). This situation is comparable with a *t* test in which a trivial difference becomes significant solely because of sample size (Friston, 2012). We followed the recommendation to have approximately 20 people per item (Linacre, 1994). Second, we fit a Rasch model in each subgroup and examined the mean-square infit statistic of each item. When the response pattern for a symptom fits the Rasch model, this statistic is close to 1 (Wright & Linacre, 1994). Each item’s mean-square infit value should be within the interval 0.89 to 1.11 given our sample size (Smith, Schumacker, & Bush, 1998). Items with infit mean squares outside this range were dropped. Third, we examined local independence. The Rasch model postulates that responses to the items should be uncorrelated after accounting for the latent trait. In practice, this meant that residual correlations be smaller than .2 and *p* values (adjusted for multiple comparisons using Holm’s method) be larger than .05. Fourth, differential item functioning (DIF) analysis was performed. The rationale for DIF analysis is to develop a measure that is invariant with respect to irrelevant characteristics such as gender or age. We tested

for both uniform DIF (in which one group is more likely to endorse symptoms at all levels of emotion dysregulation) and nonuniform DIF (in which the difference in endorsement rates varies by level of emotional dysregulation). We used the Mantel–Haenszel and Breslow–Day tests to detect uniform and nonuniform DIF, respectively. An item with DIF is indicated by a significant p value of either test, following Penfield and Algina’s combined decision rule (Penfield & Algina, 2003). We set each individual item’s p value at .002 to account for multiple comparisons. Fifth, we verified that our final set of items had satisfactory reliability using person separation index (PSI) and Cronbach’s alpha as criteria. A PSI of .7 or greater indicates that the scale reliably distinguishes two groups of participants—that is, children with and without emotional regulation.

External Validation and Optimal Threshold

We compared the resulting scale with a validated psychometric instrument, the Conners’ Global Index for Emotional Lability (EL; Conners, 2014). Conners’ EL was our chosen criterion for two reasons. First, it provides normative data for emotional lability that takes into account the participant’s age, sex, and rater. Second, our emotional dysregulation scale shares common items with EL: easily frustrated, mood changes quickly, and temper outbursts. We used area under the receiver operating characteristic (ROC) curve analysis as a concordance measure, with EL as the criterion. For this purpose, we dichotomized EL into severe emotional lability (EL percentiles at 80 and above) and nonsevere emotional lability (EL percentiles < 80), as suggested in the Conners’ testing manual (Conners, 2014). We were interested in both the global performance of our scale and the particular score that maximizes accuracy. Global performance can be interpreted as the probability that a random participant with severe EL will have a higher score in our emotional dysregulation scale than a random participant without severe EL (Hanley & Mcneil, 1982). We selected the optimal threshold according to Liu’s criterion—the score that maximizes the product of sensitivity and specificity (Liu, 2012). In this last step, we pooled together our calibration and validation samples ($n = 665$). Fifty-five observations were dropped because they were outside the age range of EL norms.

Our study received ethics approval from the university’s behavioral ethics committee. All statistical analyses were performed in R using the TAM, sirt, and difR packages (Kiefer, Robitzsch, & Wu, 2016; Magis, Beland, Tuerlinckx, & De Boeck, 2010; R Core Team, 2015; Robitzsch, 2016). Area under the ROC curve analysis was calculated using Stata.

Results

The children and youth in our study were about 9 years of age, had about a 3 to 1 male-to-female and parent-to-teacher

Table 1. Comparison of Child and Youth Samples Used for Calibration and Validation.

Characteristic	Calibration set $n = 360$	Validation set $n = 360$	p
Mean age (SD)	9.53 (3.02)	9.23 (3.13)	.20
Sex			
Female (%)	95 (26.39)	92 (25.56)	.80
Male (%)	265 (73.61)	268 (74.44)	
Rater			
Teacher	69 (19.17)	79 (21.94)	.36
Parent	291 (80.83)	281 (78.06)	
Age group			
Children (5-12 years)	298 (82.78)	294 (81.67)	.70
Youth (13-17 years)	62 (17.22)	66 (18.33)	
Mean raw score ^a	7.99 (5.21)	8.12 (5.28)	.75

^aPersons with all “yes” or all “no” answers to the 18 questions were excluded because they do not contribute to the estimation of emotion dysregulation levels or item difficulties.

ratios. Children outnumbered youth by more than 4 to 1. The mean number of items endorsed was about eight out of 18. Our calibration and validation samples did not differ materially in these variables. See Table 1 for the breakdown of these characteristics by sample.

Our analysis showed that eight of the 18 items deviated significantly from a probabilistic Guttman pattern in the calibration sample. These items were “deliberately annoys others”; “angry and resentful”; “quarrelsome”; “negative, defiant, disobedient, or hostile to authority”; “acts smart”; “excessive emotionality and attention seeking”; “instability in relationships, reactive mood, impulsivity”; and “irritable, angry outbursts, difficulty concentrating.” When the analysis was repeated in the validation sample, the results were generally consistent, except for “blames others” and “quarrelsome,” which did not misfit. See Table 2. The eight misfitting items were eliminated.

The remaining 10 items that satisfied a probabilistic Guttman pattern were then tested for local dependency. “Actively defies adult requests” had a greater than expected correlation with “argues with adults,” $\chi^2 = 10.76$, and with “uncooperative,” $\chi^2 = 9.49$, and was, therefore, eliminated. None of the remaining items violated local independence. When the analysis was repeated in the validation set, the largest residual correlation also involved “argues with adults” and “actively defies,” $\chi^2 = 18.06$. After eliminating “actively defies,” all residual correlations were less than .1 in magnitude. No item was identified as having uniform or nonuniform differentially functioning across respondent sex, rater type, or age category in the calibration sample. The same result was found in the validation sample. See Tables 3 to 5.

Questions that combined multiple concepts (i.e., “negative, defiant, disobedient, or hostile toward authority figures”;

Table 2. Item Difficulties (in Logits) and Infit Mean Squares in the Calibration and Validation Data Sets From a Set of 18 Indicators of Emotional Dysregulation.

SNAP item number		Calibration set		Validation set	
		Item difficulties	Infit M^2	Item difficulties	Infit M^2
21.	Loses temper	-0.32	0.94	-0.26	0.92
22.	Argues with adults	-0.60	0.94	-0.68	0.93
23.	Active defiance ^a	0.03	0.94	-0.23	0.92
24.	Annoys others ^b	0.24	1.15	0.21	0.97
25.	Blames others	-0.62	0.96	-0.56	1.10
26.	Touchy	-0.45	0.95	-0.33	1.08
27.	Angry, resentful ^b	0.72	0.80	0.85	0.77
28.	Spiteful	2.24	0.93	1.73	0.89
29.	Quarrelsome ^b	0.75	0.83	0.44	0.89
30.	Negative ^b	1.10	0.87	1.10	0.82
34.	Uncooperative	0.43	0.96	0.55	0.99
35.	Acts smart ^b	0.24	1.30	0.39	1.26
38.	Mood changes quickly	0.22	0.97	-0.01	1.04
39.	Easily frustrated	-0.49	0.93	-0.63	1.04
54.	Irritable	0.63	0.91	0.74	0.97
58.	Excessive emotion and attention seeking ^b	0.43	1.11	0.27	1.15
60.	Unstable relationships ^b	0.97	1.26	1.18	1.00
78.	Irritable, angry outbursts, difficulty concentrating ^b	0.51	1.16	0.36	1.15

Note. SNAP refers to Swanson Nolan and Pelham (SNAP-IV).

^aThese were eliminated due to local dependence.

^bThese were eliminated because of poor fit to the Rasch model.

“emotional, seeks attention”; “instability in relationships, reactive mood, and impulsivity”; and “irritable, angry, or difficulty concentrating”) were detected by Rasch analysis as problematic items. By contrast, statements with simple concepts (e.g., “uncooperative”) were less problematic. It is possible that multiple concept items are confusing to raters, leading to Rasch model violations.

The nine remaining items, which constitute the Clinical Evaluation of Emotional Regulation-9 (CEER-9), had satisfactory PSIs of 0.72 in the calibration sample and 0.70 in the validation sample, indicating that the scale adequately distinguishes a group of children with and without emotional dysregulation. The Cronbach's alphas of the nine retained items were .83 and .80 in the calibration and validation samples, respectively. Using the Conners' EL as the criterion, the area under the ROC curve for our nine-item scale was 0.87. The threshold score that maximized accuracy was 4. The area under the ROC curve for a score of 4 is 0.81. See Table 6 for the sensitivity and specificity at each of the 10 cut points.

To further compare the performance of the Conners' EL and CEER-9, we calculated the rates of emotional dysregulation in children and youth with ADHD. Rates of emotional dysregulation in ADHD inattentive, hyperactive, and combined types were 63%, 79%, and 88%, using the Conners' EL as compared with 42%, 56%, and 71% using CEER-9.

Discussion

Using a calibration and validation sample from a large data set of clinically referred children and youth, nine items satisfied Rasch model requirements, had adequate reliability, and was concordant with an external criterion. This subset of items total score reflects the level of emotional regulation. Having satisfied the Rasch model, there is no need for weighting the symptoms. As such, unlike previous scales, such as the Conners', it is simple to score, and it does not require the use of a table that is separated by age, rater, and gender. The present study showed that emotional regulation in children and youth can be measured using nine items derived from the SNAP-90 scale, called the CEER-9.

It is striking that the CEER-9 retains five items from ODD, the two core symptoms of DMDD, and two symptoms from the Conners' emotional lability index. Please refer to Table 7. *Uncooperative* and *easily frustrated* do not appear in the other three measures. Although *DSM-5* (APA, 2013) recommends that children who meet both the criteria for ODD and DMDD be given the single DMDD diagnosis, the present work seems to provide support for previous experts who questioned whether ODD and DMDD require separate categories (Lochman et al., 2015; Meyers et al., 2017). The CEER-9 would allow diagnosis to move toward a more empirically based grouping of symptoms that possibly stems from the same underlying process (Lochman et al., 2015).

The grouping of items in the CEER-9 suggests that externalized emotions, mood swings, being susceptible to perceived annoyances, and subtle noncooperation are all facets of emotional dysregulation. In this regard, the CEER list of items is different from the PROMIS Pediatric Anger Scale, which is made up exclusively of externalized anger. That *mood swings* is also part of the final scale suggests that the boundaries between ADHD, ODD, DMDD, and child/youth bipolar disorder categories are porous. The overlap in symptoms possibly reflects the shared underlying neural circuitry across disorders (Brotman et al., 2010; P. Shaw, Stringaris, Nigg, & Leibenluft, 2014). Accordingly, more severe emotional lability is associated with more severe presentations of ADHD and ODD (Sobanski et al., 2010).

As with all studies, the present one has several limitations. Parent- and teacher-rated scales assume that they are accurate judges of child dispositions and behaviors. Outward displays such as “losing temper” are more easily

Table 3. Uniform and Nonuniform DIF by Age Group: Children vs. Youth.

SNAP item number	Sample 1 (n = 360)					Sample 2 (n = 360)				
	Mantel–Haenszel χ^2	p	Breslow–Day	p	Combined decision rule ^a	Mantel–Haenszel χ^2	p	Breslow–Day	p	Combined decision rule ^a
21. Loses temper	0.000	.990	6.546	.478	DIF not indicated	0.000	.989	8.880	.181	DIF not indicated
22. Argues with adults	9.097	.003	8.037	.329	DIF not indicated	0.147	.701	7.299	.398	DIF not indicated
25. Blames others	0.001	.970	8.916	.178	DIF not indicated	0.014	.907	6.966	.432	DIF not indicated
26. Touchy	0.198	.657	8.177	.317	DIF not indicated	1.975	.160	8.018	.331	DIF not indicated
28. Spiteful	0.149	.699	8.832	.065	DIF not indicated	2.167	.141	4.056	.773	DIF not indicated
34. Uncooperative	4.696	.030	5.434	.607	DIF not indicated	0.371	.542	11.352	.124	DIF not indicated
38. Mood changes quickly	0.127	.721	6.446	.489	DIF not indicated	0.003	.954	5.559	.592	DIF not indicated
39. Easily frustrated	3.040	.081	14.427	.044	DIF not indicated	0.143	.706	2.844	.899	DIF not indicated
54. Irritable	0.745	.388	4.453	.616	DIF not indicated	0.601	.438	4.276	.748	DIF not indicated

Note. DIF = differential item functioning; SNAP refers to Swanson Nolan and Pelham (SNAP- IV).

^aCombined decision rule: DIF items are indicated by a p value $\leq .002$ for either the Mantel–Haenszel or Breslow–Day test.

Table 4. DIF by Child Gender.

SNAP item number	Sample 1 (n = 360)					Sample 2 (n = 360)				
	Mantel–Haenszel χ^2	p	Breslow–Day	p	Combined decision rule ^a	Mantel–Haenszel χ^2	p	Breslow–Day	p	Combined decision rule ^a
21. Loses temper	0.519	.471	5.676	.578	DIF not indicated	0.023	.881	0.913	.989	DIF not indicated
22. Argues with adults	0.079	.778	4.433	.729	DIF not indicated	0.001	.976	5.685	.577	DIF not indicated
25. Blames others	1.254	.263	7.691	.262	DIF not indicated	0.026	.873	6.521	.480	DIF not indicated
26. Touchy	0.290	.590	11.849	.106	DIF not indicated	0.002	.969	7.825	.348	DIF not indicated
28. Spiteful	0.011	.916	2.860	.582	DIF not indicated	1.100	.294	6.838	.446	DIF not indicated
34. Uncooperative	0.149	.699	3.922	.789	DIF not indicated	0.667	.414	2.718	.910	DIF not indicated
38. Mood changes quickly	4.162	.041	3.801	.802	DIF not indicated	3.816	.051	7.485	.380	DIF not indicated
39. Easily frustrated	1.180	.277	6.922	.437	DIF not indicated	0.361	.548	14.982	.036	DIF not indicated
54. Irritable	2.741	.098	10.859	.093	DIF not indicated	0.101	.751	5.964	.544	DIF not indicated

Note. DIF = differential item functioning; SNAP refers to Swanson Nolan and Pelham (SNAP- IV).

^aCombined decision rule: DIF items are indicated by a p value $\leq .002$ for either the Mantel–Haenszel or Breslow–Day test.

observable but “quick changes in mood” and “being spiteful” are subjective. The SNAP-IV 90-item rating scale has parent-, teacher-, and youth-rated versions. Due to a low

number of youth respondents, we were unable to include youth-rated SNAPs in our study. It would be important to examine whether the same SNAP items composing the

Table 5. DIF by Reporter: Parent vs. Teacher.

SNAP item number	Sample 1 (n = 360)					Sample 2 (n = 360)				
	Mantel-Haenszel χ^2	p	Breslow-Day	p	Combined decision rule ^a	Mantel-Haenszel χ^2	p	Breslow-Day	p	Combined decision rule ^a
21. Loses temper	8.739	.003	5.676	.578	DIF not indicated	2.503	.114	0.913	.989	DIF not indicated
22. Argues with adults	2.717	.099	4.433	.729	DIF not indicated	0.006	.937	5.685	.577	DIF not indicated
25. Blames others	0.027	.868	7.691	.262	DIF not indicated	0.123	.726	6.521	.480	DIF not indicated
26. Touchy	1.124	.289	11.849	.106	DIF not indicated	7.947	.005	7.825	.348	DIF not indicated
28. Spiteful	0.001	.974	2.860	.582	DIF not indicated	2.073	.150	6.838	.446	DIF not indicated
34. Uncooperative	0.874	.350	3.922	.789	DIF not indicated	1.120	.290	2.718	.910	DIF not indicated
38. Mood changes quickly	0.003	.960	3.801	.802	DIF not indicated	1.747	.186	7.485	.380	DIF not indicated
39. Easily frustrated	3.501	.061	6.922	.437	DIF not indicated	0.018	.892	14.982	.036	DIF not indicated
54. Irritable	0.038	.846	10.859	.093	DIF not indicated	1.697	.193	5.964	.544	DIF not indicated

Note. DIF = differential item functioning; SNAP refers to Swanson Nolan and Pelham (SNAP-IV).
^aCombined decision rule: DIF items are indicated by a p value \leq .002 for either the Mantel-Haenszel or Breslow-Day test.

Table 6. Classification Accuracy of the Emotional Dysregulation Scale at Various Cut Points With Conners' EL 80th Percentile as the Criterion.

Cut point	Sensitivity	Specificity	Sensitivity \times Specificity
≥ 0	1.000	0.000	0.00
≥ 1	0.981	0.213	0.21
≥ 2	0.880	0.582	0.51
≥ 3	0.800	0.780	0.62
≥ 4	0.712	0.908	0.65
≥ 5	0.580	0.943	0.55
≥ 6	0.466	0.986	0.46
≥ 7	0.361	1.000	0.36
≥ 8	0.235	1.000	0.23
≥ 9	0.095	1.000	0.10

Note. EL = emotional lability. The row appearing in bold format represents the cut point that maximizes classification accuracy.

Table 7. List of CEER-9 Symptoms Shared With ODD, DMDD, and Conners' EL.

CEER-9 item	ODD	DMDD	Conners' emotional lability index
Loses temper	x	x	X
Argues with adults	x		
Blames others for mistakes	x		
Touchy or easily annoyed	x		
Spiteful or vindictive	x		
Uncooperative			
Mood changes quickly			X
Easily frustrated			
Irritable		x	

Note. CEER-9 = Clinical Evaluation of Emotional Regulation-9; ODD = oppositional defiant disorder; DMDD = disruptive mood dysregulation disorder; EL = emotional lability.

CEER are a valid measure of emotional regulation in youth. Finally, our list of symptoms and their wording were taken from the SNAP-90. It is likely that there are other symptoms or alternate wordings that serve the same purpose.

The development of a reliable, valid, evidence-based scale dedicated to measuring emotional dysregulation, unaffected by characteristics such as age and gender, and rater, is needed for clinical work and research. CEER-9 serves as a tool for quantifying the prevalence of emotional dysregulation in children and youth. Recently, irritability

scales have been developed and validated for women (Born, Koren, Lin, & Steiner, 2008) and adults (Craig, Hietanen, Markova, & Berrios, 2008), including one using IRT (Holtzman, O'Connor, Barata, & Stewart, 2015). CEER-9 will allow researchers to better understand the developmental course of emotional regulation. Because the experience and expression of emotions vary across the life span, it would be fruitful for future work to examine whether the questions in CEER might apply to older adults as well,

including those with neurocognitive disorders. Used in the context of prospective follow-up studies, the CEER-9 could be used to analyze whether adult psychiatric conditions such as, but not limited to, mood and anxiety disorders might have their origins in childhood irritability. The CEER-9 could also be used as a developmental milestone, enabling the study of genetic and environmental precursors of childhood emotional dysregulation. When used together with neuroimaging data, CEER-9 could help establish the structural and functional bases of emotional dysregulation. Finally, having a reliable and valid instrument enables the study of pre- and post- measurements of patient response to pharmacological and nonpharmacological treatments.

Conclusion

This study reports a nine-item rating scale, the CEER-9, which can be used as an observer-reported rating scale developed in children and youth, whose sum total is a measure of emotional regulation, with a score of 4 or more out of 9 indicating current emotional dysregulation. This scale has good psychometric properties, performing similarly by child sex, age group, and parent or teacher reporter, and has satisfactory PSI and good internal and external validity. As emotional dysregulation is common, present across the life span, associated with many psychiatric and medical disorders, and independently contributes to significant morbidity and mortality, the CEER-9 would be valuable in clinical practice and research applications in many areas of psychiatry and health care.

Acknowledgment

Special thanks to Dr. Rudy Bowen for his thoughtful comments on an earlier version of this work, Professor Alan Tennant for providing guidance in Rasch Analysis, and the children and adolescents, their parents, and teachers for completing the rating scales.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Dr. Quinn is a consultant for Eli-Lilly, Shire, Janssen, Purdue, and Highland Therapeutics. Dr. Duncan is on the advisory boards and speaker bureaus of Shire, Janssen, and Purdue.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.
- Andrich, D. (1988). *Rasch models for measurement*. Newbury Park, CA: SAGE.
- Barkley, R. A., & Benton, C. M. (1998). *Your defiant child: 8 steps to better behavior*. New York, NY: Guilford Press.
- Barkley, R. A., & Fischer, M. (2010). The unique contribution of emotional impulsiveness to impairment in major life activities in hyperactive children as adults. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*, 503-513. doi:10.1016/j.jaac.2010.01.019
- Barry, T. D., Marcus, D. K., Barry, C. T., & Coccaro, E. F. (2013). The latent structure of oppositional defiant disorder in children and adults. *Journal of Psychiatric Research*, *47*, 1932-1939. doi:10.1016/j.jpsychires.2013.08.016
- Birmaher, B. (2016). The risks of persistent irritability. *Journal of the American Academy of Child & Adolescent Psychiatry*, *55*, 538-539. doi:10.1016/j.jaac.2016.04.015
- Born, L., Koren, G., Lin, E., & Steiner, M. (2008). A new, female-specific irritability rating scale. *Journal of Psychiatry & Neuroscience*, *33*, 344-354. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/18592028>
- Brotman, M. A., Rich, B. A., Guyer, A. E., Lunsford, J. R., Horsey, S. E., Reising, M. M., . . . Leibenluft, E. (2010). Amygdala activation during emotion processing of neutral faces in children with severe mood dysregulation versus ADHD or bipolar disorder. *The American Journal of Psychiatry*, *167*, 61-69. doi:10.1176/appi.ajp.2009.09010043
- Brotman, M. A., Schmajuk, M., Rich, B. A., Dickstein, D. P., Guyer, A. E., Costello, E. J., . . . Leibenluft, E. (2006). Prevalence, clinical correlates, and longitudinal course of severe mood dysregulation in children. *Biological Psychiatry*, *60*, 991-997. doi:10.1016/j.biopsych.2006.08.042
- Caprara, G. V., Cinanni, V., Dimperio, G., Passerini, S., Renzi, P., & Travaglia, G. (1985). Indicators of impulsive aggression: Present status of research on irritability and emotional susceptibility scales. *Personality and Individual Differences*, *6*, 665-674. doi:10.1016/0191-8869(85)90077-7
- Caron, C., & Rutter, M. (1991). Comorbidity in child psychopathology: Concepts, issues and research strategies. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *32*, 1063-1080.
- Caspi, A., Wright, B. R. E., Moffitt, T. E., & Silva, P. A. (1998). Early failure in the labor market: Childhood and adolescent predictors of unemployment in the transition to adulthood. *American Sociological Review*, *63*, 424-451. doi:10.2307/2657557
- Cavanagh, M., Quinn, D., Duncan, D., Graham, T., & Balbuena, L. (2014). Oppositional defiant disorder is better conceptualized as a disorder of emotional regulation. *Journal of Attention Disorders*, *21*, 381-389. doi:10.1177/1087054713520221
- Cipriani, G., Vedovello, M., Nuti, A., & Di Fiorino, M. (2011). Aggressive behavior in patients with dementia: Correlates and management. *Geriatrics & Gerontology International*, *11*, 408-413. doi:10.1111/j.1447-0594.2011.00730.x
- Cole, P. M., Martin, S. E., & Dennis, T. A. (2004). Emotion regulation as a scientific construct: Methodological challenges and directions for child development research. *Child Development*, *75*, 317-333. doi:10.1111/j.1467-8624.2004.00673.x
- Cole, P. M., Michel, M. K., & Teti, L. O. (1994). The development of emotion regulation and dysregulation: A clinical perspective. *Monographs of the Society for Research in Child Development*, *59*(2-3), 73-100. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7984169>
- Conklin, C. Z., Bradley, R., & Westen, D. (2006). Affect regulation in borderline personality disorder. *The Journal of Nervous and Mental Disease*, *194*, 69-77. doi:10.1097/01.nmd.0000198138.41709.4f

- Conners, C. K. (2014). *Conners 3rd edition manual* (3rd ed.). Toronto, Ontario, Canada: Multi-Health Systems.
- Craig, K. J., Hietanen, H., Markova, I. S., & Berrios, G. E. (2008). The Irritability Questionnaire: A new scale for the measurement of irritability. *Psychiatry Research, 159*, 367-375. doi:10.1016/j.psychres.2007.03.002
- da Rocha, N. S., Chachamovich, E., de Almeida Fleck, M. P., & Tennant, A. (2013). An introduction to Rasch analysis for psychiatric practice and research. *Journal of Psychiatric Research, 47*, 141-148. doi:10.1016/j.jpsychires.2012.09.014
- Demark, J., & Gemeinhardt, M. (2002). Anger and its management for survivors of acquired brain injury. *Brain Injury, 16*, 91-108. doi:10.1080/026990501110102059
- Drabick, D. A., & Gadow, K. D. (2012). Deconstructing oppositional defiant disorder: Clinic-based evidence for an anger/irritability phenotype. *Journal of Child & Adolescent Psychiatry, 51*, 384-393. doi:10.1016/j.jaac.2012.01.010
- Fayers, P. M., & Machin, D. (2016). *Quality of life: The assessment, analysis, and reporting of patient-reported outcomes* (3rd ed.). Chichester, UK: John Wiley.
- Fergusson, D. M., John Horwood, L., & Ridder, E. M. (2005). Show me the child at seven: The consequences of conduct problems in childhood for psychosocial functioning in adulthood. *Journal of Child Psychology & Psychiatry, 46*, 837-849. doi:10.1111/j.1469-7610.2004.00387.x
- Friston, K. (2012). Ten ironic rules for non-statistical reviewers. *Neuroimage, 61*, 1300-1310. doi:10.1016/j.neuroimage.2012.04.018
- Galanter, C. A., Carlson, G. A., Jensen, P. S., Greenhill, L. L., Davies, M., Li, W., . . . Swanson, J. M. (2003). Response to methylphenidate in children with attention deficit hyperactivity disorder and manic symptoms in the multimodal treatment study of children with attention deficit hyperactivity disorder titration trial. *Journal of Child and Adolescent Psychopharmacology, 13*, 123-136. doi:10.1089/104454603322163844
- Geller, B., Zimmerman, B., Williams, M., Bolhofner, K., Craney, J. L., DelBello, M. P., . . . Soutullo, C. (2001). Reliability of the Washington University in St. Louis Kiddie Schedule for Affective Disorders and Schizophrenia (WASH-U-KSADS) mania and rapid cycling sections. *Journal of Child Psychology & Psychiatry, 40*, 450-455. doi:10.1097/00004583-200104000-00014
- Glenn, C. R., & Klonsky, E. D. (2009). Emotion dysregulation as a core feature of borderline personality disorder. *Journal of Personality Disorders, 23*, 20-28. doi:10.1521/pedi.2009.23.1.20
- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment, 26*, 41-54. doi:10.1023/B:Joba.0000007455.08539.94
- Hanley, J. A., & Mcneil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*(1), 29-36.
- Hobart, J. C., Cano, S. J., Zajicek, J. P., & Thompson, A. J. (2007). Rating scales as outcome measures for clinical trials in neurology: Problems, solutions, and recommendations. *The Lancet Neurology, 6*, 1094-1105. doi:10.1016/S1474-4422(07)70290-9
- Holtzman, S., O'Connor, B. P., Barata, P. C., & Stewart, D. E. (2015). The Brief Irritability Test (BITE): A measure of irritability for use among men and women. *Assessment, 22*, 101-115. doi:10.1177/1073191114533814
- Irwin, D. E., Stucky, B. D., Langer, M. M., Thissen, D., DeWitt, E. M., Lai, J. S., . . . DeWalt, D. A. (2012). PROMIS Pediatric Anger Scale: An item response theory analysis. *Quality of Life Research, 21*, 697-706. doi:10.1007/s11136-011-9969-5
- Kiefer, T., Robitzsch, A., & Wu, M. (2016). *TAM: Test Analysis Modules* (Version No. 1.17-0). Retrieved from <http://CRAN.R-project.org/package=TAM>
- Kim, K. Y., Moles, J. K., & Hawley, J. M. (2001). Selective serotonin reuptake inhibitors for aggressive behavior in patients with dementia after head injury. *Pharmacotherapy, 21*, 498-501. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11310524>
- Koko, K., Bergman, L. R., & Pulkkinen, L. (2003). Child personality characteristics and selection into long-term unemployment in Finnish and Swedish longitudinal samples. *International Journal of Behavioral Development, 27*, 134-144. doi:10.1080/01650250244000137
- Linacre, J. M. (1994). Samples size and item calibration stability. *Rasch Measurement Transactions, 7*(4), 328. Retrieved from <http://www.rasch.org/rmt/rmt74m.htm>
- Liu, X. H. (2012). Classification accuracy and cut point selection. *Statistics in Medicine, 31*, 2676-2686. doi:10.1002/sim.4509
- Lochman, J. E., Evans, S. C., Burke, J. D., Roberts, M. C., Fite, P. J., Reed, G. M., . . . Elena Garralda, M. (2015). An empirically based alternative to DSM-5's disruptive mood dysregulation disorder for ICD-11. *World Psychiatry, 14*, 30-33. doi:10.1002/wps.20176
- Magis, D., Beland, S., & Tuerlinckx, De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods, 42*, 847-862.
- Martel, M. M., Gremillion, M., Roberts, B., von Eye, A., & Nigg, J. T. (2010). The structure of childhood disruptive behaviors. *Psychological Assessment, 22*, 816-826. doi:10.1037/a0020975
- Martin-Lof, P. (1974). The notion of redundancy and its use as a quantitative measure of the discrepancy between a statistical hypothesis and a set of observational data. *Scandinavian Journal of Statistics, 1*, 3-18.
- Meyers, E., DeSerisy, M., & Roy, A. K. (2017). Disruptive Mood Dysregulation Disorder (DMDD): An RDoC perspective. *Journal of Attention Disorders, 216*, 117-122. doi:10.1016/j.jad.2016.08.007
- Mick, E., Spencer, T., Wozniak, J., & Biederman, J. (2005). Heterogeneity of irritability in attention-deficit/hyperactivity disorder subjects with and without mood disorders. *Biological Psychiatry, 58*, 576-582. doi:10.1016/j.biopsych.2005.05.037
- Miller, J. D., & Pilkonis, P. A. (2006). Neuroticism and affective instability: The same or different? *The American Journal of Psychiatry, 163*, 839-845. doi:10.1176/ajp.2006.163.5.839
- Moreno, C., Laje, G., Blanco, C., Jiang, H., Schmidt, A. B., & Olfson, M. (2007). National trends in the outpatient diagnosis and treatment of bipolar disorder in youth. *Archives of General Psychiatry, 64*, 1032-1039. doi:10.1001/archpsyc.64.9.1032
- Pelham, W. E., Jr., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher ratings of DSM-III-R symptoms for the

- disruptive behavior disorders. *American Academy of Child & Adolescent Psychiatry*, 31, 210-218. doi:10.1097/00004583-199203000-00006
- Penfield, R. D., & Algina, J. (2003). Applying the Liu-Agresti estimator of the cumulative common odds ratio to DIF detection in polytomous items. *Journal of Educational Measurement*, 40, 353-370. doi:10.1111/j.1745-3984.2003.tb01151.x
- Pickles, A., Aglan, A., Collishaw, S., Messer, J., Rutter, M., & Maughan, B. (2010). Predictors of suicidality across the life span: The Isle of Wight study. *Psychological Medicine*, 40, 1453-1466. doi:10.1017/S0033291709991905
- R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available from <https://www.R-project.org/>
- Rich, B. A., Schmajuk, M., Perez-Edgar, K. E., Fox, N. A., Pine, D. S., & Leibenluft, E. (2007). Different psychophysiological and behavioral responses elicited by frustration in pediatric bipolar disorder and severe mood dysregulation. *The American Journal of Psychiatry*, 164, 309-317. doi:10.1176/ajp.2007.164.2.309
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2, 313-345. doi:10.1111/j.1745-6916.2007.00047.x
- Robitzsch, A. (2016). *sirt: Supplementary item response theory models* (Version No. 1.10-0). Retrieved from <http://CRAN.R-project.org/package=sirt>
- Rosenblatt, A., & Leroi, I. (2000). Neuropsychiatry of Huntington's disease and other basal ganglia disorders. *Psychosomatics*, 41, 24-30. doi:10.1016/S0033-3182(00)71170-4
- Roy, A. K., Lopes, V., & Klein, R. G. (2014). Disruptive mood dysregulation disorder: A new diagnostic approach to chronic irritability in youth. *The American Journal of Psychiatry*, 171, 918-924. doi:10.1176/appi.ajp.2014.13101301
- Sahl, J. C., Cohen, L. H., & Dasch, K. B. (2009). Hostility, interpersonal competence, and daily dependent stress: A daily model of stress generation. *Cognitive Therapy and Research*, 33, 199-210. doi:10.1007/s10608-007-9175-5
- Shaw, D. S., Owens, E. B., Giovannelli, J., & Winslow, E. B. (2001). Infant and toddler pathways leading to early externalizing disorders. *American Academy of Child & Adolescent Psychiatry*, 40, 36-43. doi:10.1097/00004583-200101000-00014
- Shaw, P., Stringaris, A., Nigg, J., & Leibenluft, E. (2014). Emotion dysregulation in attention deficit hyperactivity disorder. *The American Journal of Psychiatry*, 171, 276-293. doi:10.1176/appi.ajp.2013.13070966
- Smith, R. M., Schumacker, R. E., & Bush, M. J. (1998). Using item mean squares to evaluate fit to the Rasch model. *Journal of Outcome Measurement*, 2(1), 66-78.
- Sobanski, E., Banaschewski, T., Asherson, P., Buitelaar, J., Chen, W., Franke, B., . . . Faraone, S. V. (2010). Emotional lability in children and adolescents with attention deficit/hyperactivity disorder (ADHD): Clinical correlates and familial prevalence. *The Journal of Child Psychology and Psychiatry*, 51, 915-923. doi:10.1111/j.1469-7610.2010.02217.x
- Stringaris, A. (2011). Irritability in children and adolescents: A challenge for DSM-5. *European Child & Adolescent Psychiatry*, 20, 61-66. doi:10.1007/s00787-010-0150-4
- Stringaris, A., Cohen, P., Pine, D. S., & Leibenluft, E. (2009). Adult outcomes of youth irritability: A 20-year prospective community-based study. *The American Journal of Psychiatry*, 166, 1048-1054. doi:10.1176/appi.ajp.2009.08121849
- Stringaris, A., Goodman, R., Ferdinando, S., Razdan, V., Muhrer, E., Leibenluft, E., & Brotman, M. A. (2012). The affective reactivity index: A concise irritability scale for clinical and research settings. *The Journal of Child Psychology and Psychiatry*, 53, 1109-1117. doi:10.1111/j.1469-7610.2012.02561.x
- Stringaris, A., Maughan, B., & Goodman, R. (2010). What's in a disruptive disorder? Temperamental antecedents of oppositional defiant disorder: Findings from the Avon longitudinal study. *American Academy of Child & Adolescent Psychiatry*, 49, 474-483. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/20431467>
- Swanson, J. M. (unpublished). *SNAP IV Teacher and Parent Rating Scale*. Retrieved from https://www.crfht.ca/files/8913/7597/8069/SNAPIV_000.pdf
- van der Velde, G., Beaton, D., Hogg-Johnston, S., Hurwitz, E., & Tennant, A. (2009). Rasch analysis provides new insights into the measurement properties of the neck disability index. *Arthritis Care & Research*, 61, 544-551. doi:10.1002/art.24399
- van Schuur, W. H. (2003). Mken Scale Analysis: Between the Guttman scale and parametric item response theory. *Political Analysis*, 11, 139-163.
- van Stralen, J. (2016). Emotional dysregulation in children with attention-deficit/hyperactivity disorder. *ADHD: Attention Deficit and Hyperactivity Disorders*, 8, 175-187. doi:10.1007/s12402-016-0199-0
- Weinberg, A., & Klonsky, E. D. (2009). Measurement of emotion dysregulation in adolescents. *Psychological Assessment*, 21, 616-621. doi:10.1037/a0016669
- Widiger, T. A. (1998). Four out of five ain't bad. *Archives of General Psychiatry*, 55(10), 865-866. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9783556>
- Winsper, C., Hall, J., Strauss, V. Y., & Wolke, D. (2017). Aetiological pathways to Borderline Personality Disorder symptoms in early adolescence: Childhood dysregulated behaviour, maladaptive parenting and bully victimisation. *Borderline Personality Disorder and Emotion Dysregulation*, 4, Article 10. doi:10.1186/s40479-017-0060-x
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370.

Author Biographies

Jenna Pylypow is a fifth-year psychiatry resident at the University of Saskatchewan and is currently completing a fellowship in child and adolescent psychiatry.

Declan Quinn is a child and adolescent psychiatrist at the University of Saskatchewan and was formerly the head of the Division of Child and Adolescent Psychiatry.

Don Duncan is a child and adolescent psychiatrist and Clinical Director of the BC Interior ADHD Clinic in Kelowna, British Columbia. He administers www.adhdratingscales.com.

Lloyd Balbuena is a researcher with the Department of Psychiatry at the University of Saskatchewan.