

The Emotion Dysregulation Inventory: Psychometric Properties and Item Response Theory Calibration in an Autism Spectrum Disorder Sample

Carla A. Mazefsky , Lan Yu, Susan W. White, Matthew Siegel, and Paul A. Pilkonis

Individuals with autism spectrum disorder (ASD) often present with prominent emotion dysregulation that requires treatment but can be difficult to measure. The Emotion Dysregulation Inventory (EDI) was created using methods developed by the Patient-Reported Outcomes Measurement Information System (PROMIS[®]) to capture observable indicators of poor emotion regulation. Caregivers of 1,755 youth with ASD completed 66 candidate EDI items, and the final 30 items were selected based on classical test theory and item response theory (IRT) analyses. The analyses identified two factors: (a) Reactivity, characterized by intense, rapidly escalating, sustained, and poorly regulated negative emotional reactions, and (b) Dysphoria, characterized by anhedonia, sadness, and nervousness. The final items did not show differential item functioning (DIF) based on gender, age, intellectual ability, or verbal ability. Because the final items were calibrated using IRT, even a small number of items offers high precision, minimizing respondent burden. IRT co-calibration of the EDI with related measures demonstrated its superiority in assessing the severity of emotion dysregulation with as few as seven items. Validity of the EDI was supported by expert review, its association with related constructs (e.g., anxiety and depression symptoms, aggression), higher scores in psychiatric inpatients with ASD compared to a community ASD sample, and demonstration of test-retest stability and sensitivity to change. In sum, the EDI provides an efficient and sensitive method to measure emotion dysregulation for clinical assessment, monitoring, and research in youth with ASD of any level of cognitive or verbal ability. *Autism Res* 2018, 11: 928–941. © 2018 International Society for Autism Research, Wiley Periodicals, Inc.

Lay Summary: This paper describes a new measure of poor emotional control called the Emotion Dysregulation Inventory (EDI). Caregivers of 1,755 youth with ASD completed candidate items, and advanced statistical techniques were applied to identify the best final items. The EDI is unique because it captures common emotional problems in ASD and is appropriate for both nonverbal and verbal youth. It is an efficient and sensitive measure for use in clinical assessments, monitoring, and research with youth with ASD.

Keywords: autism spectrum disorder; emotion regulation; irritability; assessment; questionnaire; item response theory; PROMIS[®]

Introduction

Emotion Dysregulation in ASD

Emotion dysregulation, or difficulty modulating emotion in the service of one's goals, is common in individuals with autism spectrum disorder (ASD) [Weiss, Riosa, Mazefsky, & Beaumont, 2017]. Delayed, maladaptive, or ineffective emotion regulation has been found in young children [Konstantareas & Stewart, 2006; Nuske et al., 2017] and older children, adolescents, and adults [Mazefsky, Borue, Day, & Minshew, 2014; Samson, Huber, Gross, 2012; Samson, Hardan, Lee, Phillips, & Gross, 2015] with ASD. Irritability is one of the most common

manifestations of emotion dysregulation, present to at least a moderate degree in over 80% of youth with ASD [Mayes et al., 2017]. Emotion dysregulation has also been proposed as a potential mechanism to explain the high rates of diagnosis of comorbid psychiatric disorders in ASD [Mazefsky et al., 2013; White et al., 2014]. Indeed, studies indicate that less effective emotion regulation is associated with more symptoms of depression and anxiety [Mazefsky, Borue, Day, & Minshew, 2014; Rieffe et al., 2011] and externalizing behaviors, in people with ASD [e.g., Patel, Day, Jones, & Mazefsky, 2016; Pouw, Rieffe, Stockmann, & Gadow, 2013; Ting &

From the Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania (C.M.A., P.P.A.); Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania (L.Y.); Department of Psychology, Virginia Tech, Blacksburg, Virginia (S.W.W.); Maine Medical Center Research Institute, Maine Behavioral Healthcare, Scarborough, Maine (M.S.)

Received September 20, 2017; accepted for publication March 05, 2018

Address for correspondence and reprints: Carla A. Mazefsky, Department of Psychiatry, 3811 O'Hara Street, Webster Hall Suite 300, Pittsburgh, PA 15213. E-mail: mazefskyca@upmc.edu

Published online 6 April 2018 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/aur.1947

© 2018 International Society for Autism Research, Wiley Periodicals, Inc.

Weiss, 2017; Samson, Hardan, Podell, Phillips, & Gross, 2015].

In addition to the impairment related to psychiatric and behavioral problems, cross-sectional research suggests that emotion dysregulation is associated with poor functioning in other domains. Emotion dysregulation can exacerbate social deficits and interfere with the ability to make social and communication gains if not addressed [Prizant, Wetherby, Rubin, & Laurent, 2003]. Further, personal accounts from individuals with ASD, teachers, and caregivers identify emotion dysregulation as a critical barrier that interferes with success in mainstream secondary classrooms [Ashburner, Ziviani, & Rodger, 2010] and the transition into college [White et al., 2016].

Measurement of Emotion Dysregulation in ASD

Studies of emotion dysregulation in ASD have relied on measures that were developed for typically developing populations, and their psychometric properties have not been investigated in ASD [Weiss, Thomson, & Chan, 2014]. There is evidence that psychometric properties can differ substantially in ASD, even for measures that are psychometrically robust in non-ASD samples [e.g., White et al., 2015]. Another challenge with utilizing measures developed in the general population is the wide range of verbal and cognitive abilities that characterize ASD. Most measures of emotional constructs include at least some items that would be inappropriate in a minimally verbal child [e.g., “can say when s/he is feeling sad, angry, or mad or fearful or afraid” from the *Emotion Regulation Checklist*; Shields & Cicchetti, 1997]. Self-report emotion regulation measures have been utilized with higher-functioning samples of individuals with ASD, but are not appropriate for those with significant intellectual disability. One option for assessing emotion regulation with less verbal individuals is coding responses to structured tasks [Jahromi, Meek, & Ober-Reynolds, 2012; Nuske et al., 2017; Zantinge, van Rijn, Stockmann, & Swaab, 2017], which can be informative in the research setting, but is time-intensive and therefore not practical for routine clinical care.

One caregiver report measure developed for ASD and appropriate across the full range of verbal ability is the Aberrant Behavior Checklist (ABC) [Aman, Singh, Stewart, & Field, 1985a,b]. Its Irritability Subscale (ABC-I) has been widely used in clinical trials with success, but it may be a better measure of behavioral, rather than emotional, dysregulation [Mikita et al., 2015]. Factor loadings of the ABC suggest that four of the five highest loading items on the Irritability Subscale assess self-injurious behavior and aggression [Kaat, Lecavalier, & Aman, 2014]. As such, there is a need for brief measures of emotion dysregulation validated for ASD that can be

used with youth of any cognitive or verbal ability as part of a multi-modal battery assessing various aspects of emotion regulation.

Development of the Emotion Dysregulation Inventory

To address this gap, we developed a caregiver report measure of emotion dysregulation that taps observable signs of poor emotion regulation, called the Emotion Dysregulation Inventory (EDI). We began with caregiver report by questionnaire because it is an efficient method of assessment that is appropriate in situations when the individual is unable to report due to cognitive or other limitations [Irwin et al., 2012]. Obtaining caregiver perspectives on emotion dysregulation is also helpful given that limited emotional insight and awareness is common in ASD [Griffin, Lombardo, & Auyeung, 2016]. In addition, health care utilization is predicted by maternal perceptions of child health and functioning [Janicke, Finney, & Riley, 2001; Ravindran & Myers, 2012], which suggests that caregiver perspective is an important component of assessment in clinical settings.

Determining what content to include as representative of emotion dysregulation is a complex challenge because there are many theories that emphasize different components and processes [see Coppin & Sander, 2016 and Gross, 2013 for review]. A recurring issue is the distinction between the initial experience of emotion and its regulation. Although there are theoretical reasons to distinguish them, they are often difficult to separate in practice. In fact, a recent investigation of questionnaires designed to measure either emotion regulation strategies or emotional reactivity concluded that “the tendency to experience strong negative emotion appeared indistinguishable from the failure to regulate such emotions” [Zelkowitz & Cole, 2016]. Thus, items tapping both emotional experience/reactivity and regulation were included in the EDI item bank [Mazefsky et al., 2016] and we explored the dimensionality of emotion dysregulation empirically.

The techniques we utilized to develop the EDI items and the psychometric analyses described here are based on methodologies developed for the Patient-Reported Outcomes Measurement Information System (PROMIS[®]), an NIH Roadmap initiative designed to improve self-reported outcomes using state-of-the-art psychometric methods (for detailed information, see <http://www.nihpromis.org/>). A distinction between PROMIS and traditional measure development in ASD is the emphasis on Item Response Theory (IRT) analyses. IRT offers many advantages that result in improved precision [Embretson & Reise, 2000]. In particular, IRT models provide information about how well single items, as well as the full set of items, discriminate between people with differing degrees of severity in addition to

Table 1. Sample Characteristics

	Combined (<i>n</i> = 1755)	IAN (<i>n</i> = 1,323)	AIC (<i>n</i> = 432)
ASD confirmation		Clinical diagnosis + SCQ > 12	ADOS-confirmed + expert opinion
Age	M = 12.3 (SD = 3.2) Range 4–20	M = 12.1 (SD = 3.2) Range 6–17.9	M = 13.1 (SD = 3.3) Range 4–20
Intellectual ability (<i>n</i> = 891)	27.8% (<i>n</i> = 248)	25.9% IQ < 70 (<i>n</i> = 175)	34.0% IQ < 70 (<i>n</i> = 73)
Verbally fluent (<i>n</i> = 1719)	55.3% (<i>n</i> = 951)	56.9% (<i>n</i> = 732)	50.7% (<i>n</i> = 219)
Gender (<i>n</i> = 1755)	20.9% female (<i>n</i> = 367)	20.9% female (<i>n</i> = 277)	20.8% female (<i>n</i> = 90)
Race	89.5% (<i>n</i> = 1571)	91.5% (<i>n</i> = 1210)	83.6% (<i>n</i> = 361)
White			
African-American	7.2% (<i>n</i> = 127)	6.2% (<i>n</i> = 82)	10.4% (<i>n</i> = 45)
Ethnicity-Hispanic	8.3% (<i>n</i> = 146)	9.6% (<i>n</i> = 127)	95.3% (<i>n</i> = 19)

Note: Intellectual ability was estimated based on direct assessment with the nonverbal intelligence quotient (NVIQ) standard score of the *Leiter International Performance Scale—Third Edition* in the AIC, and by caregiver report to the question “what was (child’s name) most recent IQ score?” with response options in ranges in IAN. Participants were considered verbally fluent if they required an ADOS module 3 or 4 for the AIC, consistent with standard ADOS administration guidelines, or were rated by caregivers as having “meaningful, fluent speech” for IAN.

providing methods to assess whether items function differently based on certain demographic characteristics (e.g., gender or verbal ability).

Following the guidelines set forth by PROMIS, the objectives of this study were to use a large ASD sample to: (1) determine the dimensionality (e.g., factor structure) of emotion dysregulation as measured by EDI items, (2) evaluate the psychometric properties of individual EDI items to identify the most sensitive and psychometrically robust items; (3) ensure that the EDI items are not biased within the ASD population by identifying any DIF based on gender, age, intellectual ability and verbal ability, and (4) examine the validity of the EDI by (a) assessing the convergence between the EDI and measures of other related constructs, (b) establishing criterion validity by comparing groups with expected mean differences (e.g., a community ASD sample and a sample of psychiatric inpatients with ASD), and (c) investigating the EDI’s test-retest stability and sensitivity to change.

Methods

Overview

The item development process was described in detail in Mazefsky et al. [2016]. In brief, the item pool was generated based on a comprehensive literature review, generation of a conceptual model, and assignment of items to an item hierarchy to ensure adequate coverage of key constructs. Once the initial item pool was developed, interviews were completed with 19 parents of youth with ASD to assess their understanding of the items and their decision-making processes when selecting their responses. Information generated from these interviews, along with input from a panel of experts in measure development and emotion dysregulation in ASD, was utilized to revise the items, directions, and response options and arrive at the final 66 candidate

items that were used for psychometric analyses and calibration.

Participants

Our sampling strategy utilized two sources to obtain a sample that was representative of the full spectrum of severity of ASD (The Interactive Autism Network; IAN) while also being enriched with the most extreme forms of emotion dysregulation in ASD (Autism Inpatient Collection; AIC) (see Table 1). Overall, the sample was predominantly Caucasian and non-Hispanic.

The Interactive Autism Network. IAN is an online registry of individuals with parent-reported professional ASD diagnoses in the United States that was developed to support internet-based research studies and aid in recruitment. Participants in IAN’s registry were invited to complete this study if they had a *Social Communication Score-Lifetime Version* (SCQ; Rutter, Bailey, and Lord, 2003) of ≥ 12 and were between the ages of 6–17 years old. Parent-reported professional diagnosis of ASD has been verified by medical records [Daniels et al., 2012]. Community professional diagnosis of ASD has also been validated [Lee et al., 2010; Marvin et al., 2014]. Invitations to participate in this study were sent to 11,648 registrants, 9,926 did not respond, 1,642 expressed an interest, and 1,323 participated.

The Autism Inpatient Collection. The AIC is a six-site study of children, adolescents, and young adults admitted to specialized inpatient psychiatric units for youth with ASD and other developmental disorders. The full methods of the AIC have been published [Siegel et al., 2015]. The AIC included patients between the ages of 4–20 years old, though very few were younger than 6 (*n* = 9). Participants with a score of ≥ 12 on the SCQ or high suspicion of ASD from the inpatient

clinical treatment team were eligible for enrollment. Inclusion in the AIC dataset required confirmation of ASD diagnosis by research-reliable administration of the *Autism Diagnostic Observation Schedule-2* [ADOS-2; Lord et al., 2000]. Exclusion criteria were the lack of availability of a caregiver proficient in English or status as a prisoner for the individual with ASD.

Measures

Emotion Dysregulation Inventory. The EDI item bank for psychometric analysis consisted of 66 items (Mazefsky et al., 2016) rated on a five-point scale of problem severity over the past 7 days: Not at all = 0, Mild = 1, Moderate = 2, Severe = 3, Very Severe = 4.

Aberrant Behavior Checklist. The ABC is a caregiver report of problem behavior for children and adults with developmental disabilities [Aman et al., 1985a, 1985b]. It has five subscales, with items scored on a 7-point Likert scale: (I) Irritability (15 items; tapping various aspects of behavioral dysregulation), (II) Lethargy/Social Withdrawal (16 items), (III) Stereotypic Behavior (7 items), (IV) Hyperactivity (16 items), and (V) Inappropriate Speech (4 items).

Child Behavior Checklist. The Child Behavior Checklist (CBCL) is a widely used measure of psychiatric symptoms [Achenbach & Rescorla, 2001]. Items are scored in reference to the past month as 0 = not true, 1 = sometimes or somewhat true, and 2 = very true or often true, producing scores for several specific syndrome scales. The scales used in analyses included: Anxious/Depressed (13 items), Withdrawn/Depressed (8 items), and Aggressive Behavior (18 items). In addition, Samson et al [2014] derived an Emotion Dysregulation Index based on 18 CBCL items that experts in emotion regulation agreed captured the construct. Because Samson et al. [2014] completed this work in an ASD sample, it was also used in analyses of convergent validity as a supplement to the CBCL's standardized scales.

Procedures

All participants were first determined to meet inclusion criteria regarding ASD status and age. IAN participants were asked to complete the EDI and ABC concurrently. Those who completed this step ($n = 1,435$) were then asked to complete additional questionnaires including the CBCL. IAN participants completed the EDI a second time 4 weeks later, together with a questionnaire that asked about any changes in treatment during that time period. For the AIC, the EDI and ABC are part of the core battery completed by caregivers during the first week of their child's inpatient stay. The CBCL was

completed during the stay. The CBCL was added to the AIC protocol in Year 3 of data collection, so CBCL data were not available for all inpatients. The EDI was completed by caregivers a second time at discharge.

Psychometric Analysis

Factor analysis. We did not expect that all 66 EDI items would reflect a single underlying trait. Although item development was informed by a conceptual model, our goal was to identify the most robust latent constructs empirically and to document sufficient unidimensionality for each of them so that we could proceed with IRT analyses in which the credibility of model parameters relies on the assumption of unidimensionality. In addition, we wanted to explore if the factor structures were the same for the IAN sample and the combined (IAN + AIC) sample. Therefore, the IAN sample ($n = 1323$) and the combined sample ($n = 1755$) were randomly split into two, non-overlapping subsamples: One for exploratory factor analysis (EFA; n for IAN = 680, n for combined = 870) and the other for confirmatory factor analysis (CFA: n for IAN = 643, n for combined = 885). Both EFA and CFA were conducted using Mplus 6.2 with promax rotation [Muthen & Muthen, 2007]. Factor loadings, scree plots, and eigenvalues were evaluated. We focused on the ratio of eigenvalues in EFAs and the relative proportions of variance accounted for by the factors extracted. We also emphasized the magnitude of factor loadings that appeared in both EFAs and CFAs and the fit and information values reflected in IRT models. If the factor structures were the same, we planned to use the combined sample for the IRT analysis.

IRT analysis. The most commonly used IRT model for polytomous items (i.e., items with 3 or more ordinal response categories) is the two-parameter graded response model [GRM; Samejima, 1969]. The GRM has a slope parameter and $n-1$ threshold parameters for each item, where n is the number of response categories. The slope parameter measures item discrimination, i.e., how well the item differentiates higher versus lower levels of severity (or Θ in IRT terms). Useful items have larger slope parameters. Threshold parameters measure item difficulty, i.e., the ease versus difficulty of endorsing different response options for an item. For example, the first threshold parameter for an item tells us where along the Θ scale of severity a respondent is more likely to endorse a response of "mild" rather than "not at all" (see Fig. 1 for an example from the EDI).

Items remaining in the pool for each construct, i.e. factor, were calibrated with the two-parameter GRM using IRTPRO 3.1. The convergence criterion for the EM cycles was set to 0.0001, with the number of cycles set to 100. Summed χ^2 were used to evaluate the model

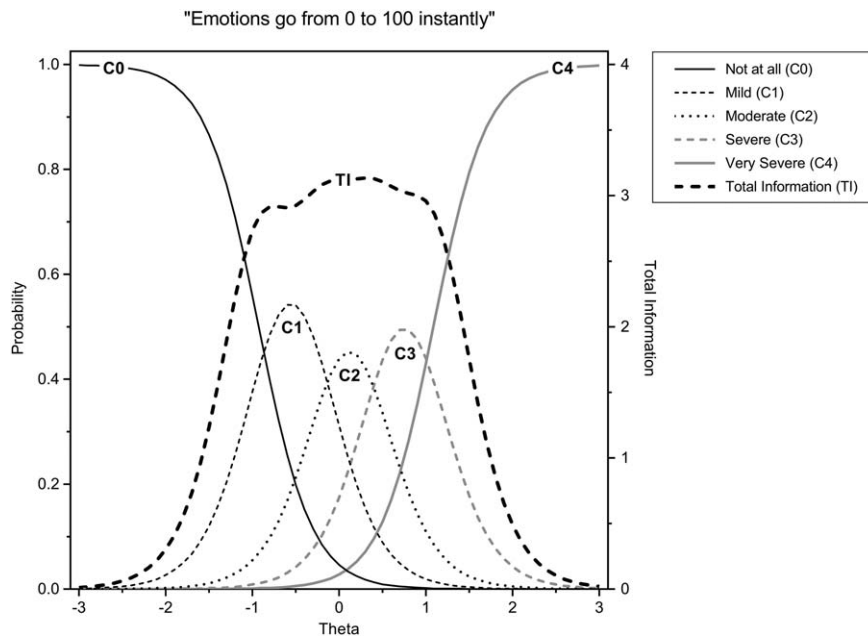


Figure 1. Example of theta and total information scores from an item on the EDI.

fit at the item level. Each item pair within each factor was evaluated for local dependency using LD χ^2 .

Differential item functioning analysis. DIF occurs when characteristics such as gender, which may seem extraneous to the assessment of the constructs under consideration, actually do have an effect on measurement. An item is flagged for DIF if it is more (or less) difficult to endorse or more (or less) discriminating in some focal group (compared to a reference group) when the different subgroups have been matched on the latent trait under investigation. We conducted DIF analyses (for both uniform and non-uniform DIF) on the basis of gender, age, verbal ability, and intellectual ability. For verbal ability, participants were considered fluent if they required an ADOS module 3 or 4 for the AIC, consistent with standard ADOS administration guidelines, or were rated as having “meaningful, fluent speech” in response to the question “how would you best describe (child name’s) current verbal ability” for IAN. Intellectual ability was estimated based on direct assessment with the nonverbal intelligence quotient (NVIQ) standard score of the *Leiter International Performance Scale—Third Edition* [Leiter-3; Roid, Miller, Pomplun, and Koch, 2013] in the AIC, and by caregiver report to the question “what was (child’s name) most recent IQ score?” with response options in ranges in IAN. To dichotomize for DIF, IQ was categorized as ≥ 70 or < 70 . Two different DIF procedures were employed: the IRT likelihood ratio method [Thissen, Steinberg, & Wainer, 1993] embedded in IRTPRO and an ordinal logistic regression procedure [Zumbo, 1999]. Items were

considered for removal if they showed significant DIF ($P < 0.01$) by both methods [Teresi et al., 2009].

Concurrent calibrations with ABC and CBCL. Concurrent calibration refers to estimating item parameters across multiple measures on a single computer run. We fixed the final item parameters for the EDI bank and calibrated relevant ABC and the CBCL items with these same parameters using the GRM. The procedure places all items on the same latent trait scale of the EDI bank. We compared the EDI Reactivity item bank, its short form (EDI-SF), and the CBCL Emotion Dysregulation Index (CBCL-EDI) as another index of emotion dysregulation that has been used in ASD, with commonly used measures of related constructs, including the ABC-I and CBCL Anxious/Depressed scale.

Convergent and criterion validity and change sensitivity evidence. To evaluate convergent and discriminant validity, we correlated theta scores on the EDI domains with ABC subscales and CBCL subscales by IAN sample, AIC sample, and the combined EDI sample. We compared means between groups expected to differ (more emotion dysregulation expected in the inpatient versus community sample). To evaluate test–retest stability and change-sensitivity, paired t -tests were utilized to compare 4-week retest theta scores in IAN participants whose caregivers reported no changes in services or treatments during that time span (scores expected to be similar at both time points), and between admission and discharge scores

for AIC participants (scores expected to be lower at discharge).

Results

Sample Characteristics

Table 1 summarizes the characteristics of the IAN sample, the AIC sample, and the combined sample. The age and gender distributions were similar between IAN and AIC, but AIC had significantly higher proportions of low IQ and low verbal ability.

Factor Structures in the Two Samples

Given the differences of IQ and verbal ability for the IAN and AIC samples, EFA was performed separately and with the combined sample. We compared the eigenvalues of 1- through 5-factor structures for these three samples. The eigenvalues and factor structures were similar across the three samples; therefore, the combined sample was used for all additional analysis. The 1- and 2-factor solutions emerged as the most meaningful, given the scree plots, the magnitude of eigenvalues, and clinical interpretation. Factor 1 (F1) was characterized by items capturing rapidly escalating, intense, and labile negative affect as well as difficulty down-regulating that affect (sustained reactions and trouble calming down). Factor 2 (F2) included items that reflect common definitions of general negative affect (sadness, unease, and anxiety) as well as low motivation. The second round of EFA on the combined sample retained 53 items with the same general content (36 items loaded on F1 and 17 items loaded on F2) after we deleted 13 items with smaller factor loadings (less than 0.45). No items had >0.45 loading on both factors. The correlations between the two factors was 0.51.

We performed single-factor CFAs on the reduced item pools to confirm their unidimensionality, using the second half of the sample ($n = 885$). For F1, all factor loadings were greater than .50, and several fit indices were strong (CFI = 0.96, TLI = 0.96) or adequate (RMSEA = 0.086 with 90% CI of 0.085–0.088). For F2, all factor loadings were larger than .45 and fit indices were more modest, revealing less homogeneity in CFA terms (reflected primarily in a larger RMSEA: RMSEA = 0.117 with 90% CI of 0.111–0.122, CFI = 0.86, TLI = 0.84).

IRT Calibrations

The two item banks were calibrated separately using the two-parameter GRM in IRTPRO. For F1, nine items with item information less than 1.0 and discrimination parameter less than 1.0 were removed, leaving a total of 27 items. For F2, five items with item information and discrimination parameter estimates less than or equal to 1.0 were removed, leaving a total of 12 items.

Following a second round of IRT calibration, one item was eliminated on the basis of model misfit ($P < 0.001$) from F1. We also examined local dependency (i.e., residual correlations) in the IRT models using the LD χ^2 . Two additional items from F1 and one item from F2 were eliminated due to local dependency. Finally, our analyses of DIF by gender, age, verbal ability and IQ flagged no items by both DIF methods, and no further items were eliminated for this reason. Item information curves were also examined to eliminate items with limited information, i.e., >0.50 . For F2, an additional five items were eliminated due to limited item information.

Thus, the final calibrated item banks included 30 items: 24 items for F1 and 6 items for F2. The final items had a Flesch Reading Ease score of 71.9 (on a scale of 0–100 with higher scores indicating easier to read) and a Flesch-Kincaid grade equivalent of 5.1. The final items, together with their IRT parameters, are summarized in Tables 2 and 3. In general, the discrimination parameter estimates were larger for F1 than those for F2, and the range of threshold parameter estimates was wider for F1 than F2, results attributable in part to the larger number of items in F1 that survived the process of pruning items. Test information curves and plots of corresponding standard errors are displayed in Figures 2 and 3. Information values of 10 correspond approximately to CTT reliability of .90. At this threshold, the effective range of measurement varied between F1 and F2, but in both cases, they were substantial: F1, -2 to $+2.5$ SDs, and F2, 0 to $+2.5$ SDs. Although F2 offers less precision below the average score, it measures negative affect so has the most clinical relevance when elevated. However, overall, F1 was more robust than F2, with more items, better IRT parameters, and smaller standard errors.

Table 4 shows the final structure of the EDI, including the two factors, the content of the items nested within each factor, and the number of items in each content area. Factor 1 was named “Reactivity” and Factor 2 was named “Dysphoria.” The correlation between the two EDI factors was .63 in the combined sample, .50 for AIC, and .59 for IAN, all significant at $P < 0.001$.

Selection of Items for Short Forms

To be able to provide a static short form as an alternative for computerized adaptive testing (CAT) administration, we rank ordered F1 items on four criteria: discrimination parameters, the percentage of times the item would have been selected in a simulated CAT using our calibration sample, expected information under the standard normal distribution with a mean of 0 and SD of 1, and expected information under a normal distribution with a larger SD, i.e., a mean of 0 and SD of 1.5 [Choi, Reise, Pilkonis, Hays, & Cella, 2010]. The CAT simulations were performed using the Firestar

Table 2. Item Parameter Estimates for F1/Reactivity in Descending Order of the Slope Parameter

Item	Stem	a	b1	b2	b3	b4
^a EDI3	Has explosive outbursts	3.66	-0.83	-0.2	0.43	1.12
^a EDI21	Hard to calm him/her down when mad or upset	3.53	-1.21	-0.38	0.37	1.06
^a EDI19	Has extreme or intense emotional reactions	3.43	-1.06	-0.26	0.48	1.26
^a EDI36	Has trouble calming him/herself down	3.37	-1.04	-0.15	0.56	1.32
^a EDI34	Emotions go from 0 to 100 instantly	3.26	-0.93	-0.18	0.42	1.09
^a EDI4	Cries or stays angry for 5 min or longer	3.20	-0.92	-0.17	0.55	1.30
EDI42	Seems to be in a rage	2.99	-0.08	0.46	0.99	1.61
EDI52	Has mood swings	2.98	-0.81	0.06	0.71	1.45
EDI24	Reactions are so intense that he/she has had to be removed from an activity or place	2.98	-0.48	0.04	0.60	1.14
^a EDI46	Reactions usually are more severe than the situations calls for	2.97	-1.23	-0.46	0.32	1.16
EDI1	Appears angry or irritable	2.91	-1.37	-0.33	0.65	1.55
EDI53	Difficult to distract if he/she is frustrated or upset	2.86	-1.18	-0.26	0.49	1.27
EDI59	Easily triggered/upset (you have to walk on eggshells around him/her)	2.80	-0.55	0.14	0.68	1.26
EDI6	Cannot calm down without help from someone else	2.55	-0.85	-0.02	0.80	1.72
EDI8	Frustrates easily	2.50	-1.91	-0.65	0.32	1.22
EDI28	When upset or angry, he/she stays that way for a long time	2.42	-0.54	0.28	1.08	1.80
EDI50	Becomes upset without a clear reason	2.39	-0.63	0.16	0.88	1.71
EDI26	Physically attacks people	2.36	0.00	0.50	0.96	1.51
EDI58	Cannot change his/her mood even with your best efforts	2.31	-0.51	0.35	1.10	2.02
EDI13	Breaks down (crying, screaming) if told he/she can't do something	2.23	-0.80	-0.05	0.59	1.35
EDI7	Suddenly switches to an opposite emotion (e.g. from sad to happy)	2.15	-0.62	0.25	1.04	1.91
EDI38	Tense or agitated and unable to relax	2.04	-0.71	0.23	1.11	2.00
EDI27	Seems on edge	1.95	-0.58	0.32	1.15	2.10
EDI10	Destroys property on purpose	1.82	-0.05	0.57	1.12	1.81

Note: ^adenotes short form items. Column a displays the slope parameter (how well the item discriminates between respondents with low or high reactivity). Columns b1–b4 display threshold values for individual responses (low threshold values indicate that the item is sensitive to low severity levels and high threshold values indicate that the item is sensitive to high severity levels).

Table 3. Item Parameter Estimates for F2/Dysphoria in Descending Order of the Slope Parameter

Item	Stem	a	b1	b2	b3	b4
EDI43	Very little makes him/her happy	3.89	-0.01	0.65	1.33	2.07
EDI31	Does not seem to enjoy anything	2.63	0.20	0.99	1.74	2.43
EDI63	Seems sad or unhappy	2.60	-0.49	0.56	1.27	2.14
EDI64	Appears uneasy through the day	1.93	-0.30	0.72	1.61	2.38
EDI57	Not responsive to praise or good things happening	1.41	0.20	1.26	2.42	3.48
EDI51	Refuses to leave the house or go to school or activities unless forced	1.22	-0.09	0.79	1.66	2.64

Note. Column a displays the slope parameter (how well the item discriminates between respondents with low or high dysphoria). Columns b1–b4 display threshold values for individual responses (low threshold values indicate that the item is sensitive to low severity levels and high threshold values indicate that the item is sensitive to high severity levels).

program [Choi, 2009]. For the CAT simulations, we set the minimum number of items to be administered to 8 and the maximum number of items to be administered to be the full bank. We then selected 7 items for the short form based on the convergence of the four psychometric criteria, the content of candidate items, and location parameters. In Table 2, daggers identify the items selected for the short form. The correlation between theta scores from the short form and the full bank was 0.98.

Concurrent Calibrations with ABC and CBCL

To compare the final EDI item bank and its short form to the widely used measures, items from the ABC and

CBCL were calibrated concurrently with the EDI item bank. Figure 4 displays the test information curves. Overall, the full EDI Reactivity item bank provided the most test information, in large part, because of the large number of items the full bank contains. The performance of the EDI-SF provides additional support for the EDI's utility. Even with fewer items, it provides the same amount of information as the ABC-I and more information than the CBCL-EDI and CBCL-Anxious/Depressed scale.

Convergent and Criterion Validity

Initial evidence for construct validity was based on expert review [see Mazefsky, et al. 2016]. In addition,

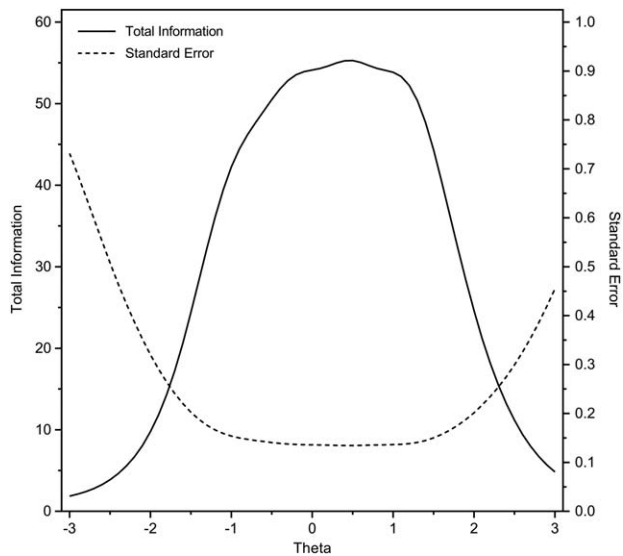


Figure 2. Total test information curves for Factor 1 (Reactivity).

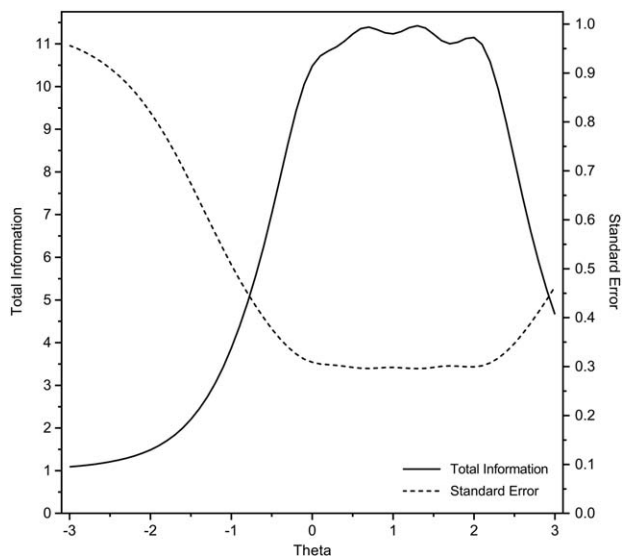


Figure 3. Total test information curves for Factor 2 (Dysphoria).

correlations with related measures revealed expected patterns (see Table 5). In particular, both EDI factors (Reactivity and Dysphoria) were positively and moderately to strongly correlated with measures of similar constructs and as expected with scores for other behavioral and emotional problems, based on prior literature demonstrating an association between emotion regulation and these constructs. Interestingly, the magnitude of correlations differed somewhat between the IAN and AIC samples, potentially due to the AIC sample including more severely affected youth and the related measures including some items that would be difficult to rate in minimally verbal or intellectually impaired youth. Finally, the AIC group had significantly higher EDI Reactivity (mean = 0.91, SD = 0.80) and EDI Dysphoria theta scores

Table 4. Content of EDI Reactivity and Dysphoria Factors

EDI Reactivity (F1)		EDI Dysphoria (F2)	
Item Content	Number of Items	Item Content	Number of Items
Sustained reactions	5	Decreased positive affect	3
Intensity	4	Nervousness and fear	2
Poor modulation	3	Increased negative affect	1
Rapid escalation	3		
Increased negative affect	3		
Lability	2		
Hyperarousal	2		
Disturbed behavior	2		

Note: Content classifications are based on the original conceptual model (Mazefsky et al., 2016). Categories are shown in order of descending number of items.

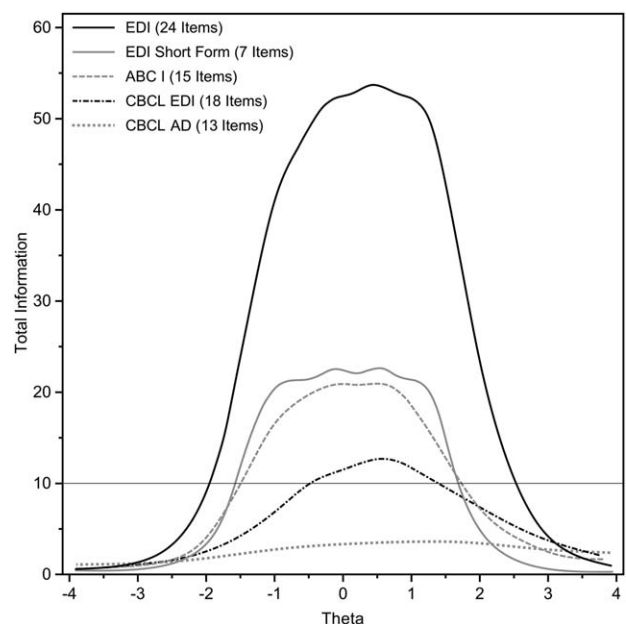


Figure 4. Test information comparison with related measures.

(mean = 0.55, SD = 0.83) than the IAN sample (EDI Reactivity: mean = -0.30, SD = 0.85, $t = 25.8$, $df = 1753$, $P < 0.001$; EDI Dysphoria: mean = -0.18, SD = 0.87, $t = 15.4$, $df = 1753$, $P < 0.001$). These findings suggest that EDI scores differ in expected ways between known groups, in support of criterion validity. In addition, the AIC inpatient group had a higher percentage of items scored as moderately intense or higher as compared to the IAN community sample (see Table 6).

Paired t -tests to evaluate the stability and change sensitivity of the Reactivity Short Form and Dysphoria theta scores are reported in Table 7. As expected, the scores were relatively stable in the IAN sample across a 4-week period, whereas there was a significant decrease in scores from admission to discharge for the AIC sample. Although there was a significant decrease in

Table 5. Correlations between Emotion Dysregulation Inventory Theta Scores and Related Measures

	Combined sample				AIC sample				IAN sample			
	Factor 1		Factor 2		Factor 1		Factor 2		Factor 1		Factor 2	
	<i>r</i>	<i>n</i>	<i>r</i>	<i>N</i>	<i>r</i>	<i>N</i>	<i>r</i>	<i>n</i>	<i>r</i>	<i>n</i>	<i>r</i>	<i>n</i>
ABC Irritability	0.835	1668	0.543	1668	0.651	418	0.347	418	0.815	1250	0.492	1250
ABC Lethargy	0.322	1696	0.540	1696	0.153	418	0.405	418	0.314	1278	0.556	1278
ABC Stereotypy	0.378	1701	0.278	1701	0.222	419	0.163	419	0.346	1282	0.241	1282
ABC Hyperactivity	0.656	1677	0.378	1677	0.473	419	0.203	419	0.606	1258	0.305	1258
ABC Speech	0.347	1707	0.190	1707	0.315	418	0.171	418	0.320	1289	0.141	1289
CBCL Samson Emotion Dysregulation Index	0.748	1263	0.575	1263	0.675	120	0.312	120	0.718	1143	0.561	1143
CBCL Anxious/Depressed	0.400	1277	0.479	1277	0.300	119	0.335	119	0.420	1158	0.496	1158
CBCL Withdrawn/Depressed	0.221	1299	0.564	1299	0.150	118	0.482	118	0.215	1181	0.573	1181
CBCL Aggressive Behavior	0.737	1252	0.482	1252	0.589	121	0.172	121	0.709	1131	0.460	1131

Note: All correlations significant at $P < .001$; ABC = Aberrant Behavior Checklist, CBCL = Child Behavior Checklist, AIC = Autism Inpatient Collection, IAN = Interactive Autism Network.

Table 6. Percent of Participants with Ratings of Moderate or Higher

Item Stem	Combined		AIC		IAN	
	N	%	n	%	n	%
Factor 1/Reactivity (Short Form)						
Reactions usually are more severe than the situation calls for	1121	63.9	370	85.6	751	56.8
Hard to calm him/her down when mad or upset	1083	61.7	388	89.8	695	52.5
Has extreme or intense emotional reactions	1015	57.8	366	84.7	649	49.1
Has explosive outbursts	969	55.2	385	89.1	584	44.1
Emotions go from 0 to 100 instantly	959	54.6	367	85.0	592	44.7
Has trouble calming him/herself down	952	54.2	372	86.1	580	43.8
Cries or stays angry for 5 min or longer	942	53.7	369	85.4	573	43.3
Factor 2/Dysphoria						
Seems sad or unhappy	559	31.9	264	61.1	295	22.3
Refuses to leave the house or go to school or activities unless forced	556	31.7	192	44.4	364	27.5
Appears uneasy throughout the day	512	29.2	236	54.6	276	20.9
Very little makes him/her happy	484	27.6	209	48.4	275	20.8
Not responsive to praise or good things happening	362	20.6	120	27.8	242	18.3
Does not seem to enjoy anything	345	19.7	118	27.3	227	17.2

Note: Items were ordered in descending order of their prevalence in the combined sample.

Reactivity scores across 4 weeks in the IAN sample, the effect size was very small, in comparison to very large effect sizes for the change in scores in the AIC sample.

Discussion

The EDI items were developed and refined through a systematic process of item development [Mazefsky et al., 2016] and psychometric evaluation with over 1,700 youth with ASD, resulting in a 24-item bank and a 7-item short form for reactivity, and a 6-item set for dysphoria. Of primary importance to the overall objective of developing a sensitive measure of observable indicators of emotion dysregulation, the test information curves indicate that the measures for reactivity and

dysphoria capture a broad range of severity with a high degree of precision. This result is especially true of the item bank and short form for reactivity, which emerged most robustly from the original item pool.

Both factor analysis and IRT results supported two separate factors rather than a single factor for emotion dysregulation. The distinction between the two factors is consistent with several theories related to the circumplex model of emotion that consider aspects such as behavioral activation [Larsen & Diener, 1992], arousal, and valence [Russell, 1980]. Specifically, Factor 1 (Reactivity) captures high arousal and negative valence characterized by irritability and anger, whereas Factor 2 (Dysphoria) captures lower arousal and negative valence characterized by general unease (sadness and anxiety). This factor structure also has some interesting parallels

Table 7. EDI Change Scores across Four Weeks in a Treatment-Stable Sample (IAN) and from Admission to Discharge in a Sample of Psychiatric Inpatients with ASD (AIC)

Theta Score-Sample	n	Score at T1 Mean (SD)	Score at T2 Mean (SD)	Difference scores Mean (SD)	Paired <i>t</i> -tests		
					<i>t</i>	<i>P</i>	Effect size
F1/Reactivity SF-IAN	901	-0.35 (0.85)	-0.40 (0.87)	0.05 (0.57)	2.94	0.003	-0.06
F1/Reactivity SF-AIC	222	0.92 (0.76)	-0.29 (0.93)	1.21 (0.97)	18.45	<0.001	-1.59
F2/Dysphoria-IAN	901	-0.21 (0.85)	-0.23 (0.86)	0.02 (0.62)	1.07	0.284	-0.03
F2/Dysphoria-AIC	222	0.50 (0.83)	-0.20 (0.80)	0.70 (0.86)	12.13	<0.001	-0.85

Note: SF = Short Form, for the IAN sample, the comparison was EDI scores completed four weeks apart. For the AIC sample, the comparison was between psychiatric hospital admission and discharge EDI scores.

to the bi-factor model of psychopathology [Cicchetti & Toth, 1991] and its internalizing (dysphoria) and externalizing (anger/reactivity) dimensions. In line with this, Factor 1 (Reactivity) was more strongly associated with measures of behavioral dysregulation (ABC Irritability, CBCL Aggression) and Factor 2 (Dysphoria) was more strongly associated with ABC Lethargy and CBCL Withdrawn/Depressed scores. It is worth emphasizing, however, that the two EDI factors were correlated and that both factors correlated with all related measures, just with some differences in magnitude.

Second, although Reactivity was considered the best overall term to describe Factor 1, it includes items that measure both initial emotional reactivity (e.g., rapidly escalating and intense reactions, negative affect) and difficulty down-regulating negative emotion once aroused (e.g., sustained emotional reactions and difficulty calming down once upset). This result is consistent with a recent factor analysis of emotion regulation and emotional reactivity measures in a typically developing population that supported a single factor capturing both intense emotionality (particularly strong negative emotion) and difficulty regulating those responses [Zelkowitz & Cole, 2016]. We note that, in temperament research, reactivity and regulation are considered separately, both conceptually and in the most commonly used assessment tools (as negative emotionality and effortful control) [Rothbart, 2006]. However, Rothbart and Sheese [2007] argue against consideration of temperament dimensions in isolation, due to the likely interactions among them. Thus, any measure of emotion dysregulation or related constructs should be part of a broader battery that considers multiple traits that may influence the presentation of problematic behavior.

Conclusions and Utility of the EDI

When using the EDI in clinical or research contexts, certain aspects of the study design should be considered. First, there may be some differences between those who chose to participate in the study and those

who did not, particularly for the sample collected online. In addition, although the AIC age minimum was 4 years old, the majority of the sample was over five years old. Therefore, the EDI is most appropriate for use with those over five. Modifications are underway to develop a version for two to five year olds. Finally, the cultural sensitivity of the EDI should be explored in large diverse samples.

Although developed to address a gap in available measures for ASD, we anticipate that the EDI is applicable outside of the ASD field as well. Emotion regulation is a transdiagnostic mechanism that is relevant to many psychiatric disorders [Aldao, Nolen-Hoeksema, & Schweitzer, 2010; Cicchetti, Ackerman, & Izard, 1995; Fernandez et al., 2016; Schäfer, Naumann, and Holmes, 2017]. The potential utility of the EDI in other populations is supported, in part, by convergence of our factor structure with the non-ASD literature [e.g., Zelkowitz & Cole, 2016]. We are currently collecting a large sample representative of the United States population to investigate the psychometric properties of the EDI outside of ASD.

The more immediate contribution of the EDI will be to understanding and supporting youth with ASD, given that it is the first measure of emotion dysregulation developed for and validated in ASD using advanced item analyses to enhance sensitivity and efficiency. Given its brevity, the EDI provides a new opportunity for screening of emotion dysregulation in ASD. It has been argued that systematic screening of emotional and behavioral concerns should be a part of routine care given the prominence of these problems among youth with ASD [Chandler et al., 2016]. The observed correlations between the EDI and aggression, withdrawal/depression, and other problematic behaviors further support the importance of identifying emotion dysregulation.

The EDI may also be useful as an outcome measure or as a mediator in treatment trials for ASD [Lerner et al., 2012]. Given that the EDI emphasizes observable signs of dysregulation rather than a single model of regulatory strategies, it can be used in comparative effectiveness studies that evaluate different treatment approaches,

as well as to determine what works best, and for whom, in line with the emphasis on personalized approaches to medicine [Norcross & Wampold, 2010]. Measures of emotion regulation strategies were not included in the EDI's psychometric battery due to the large percentage of minimally verbal participants and use of caregiver report, but determining how EDI scores vary based on patterns of emotion regulation strategy use may inform treatment development efforts.

Finally, the EDI could be utilized to inform our understanding of ASD's underlying biology. Application of the EDI with concurrent physiological assessment may be informative, particularly if the EDI factors are conceptualized as related to arousal [Russell, 1980]. In addition, there are now a handful of studies demonstrating an association between emotion regulation and neural differences in ASD [Pitskel, Bolling, Kaiser, Pelphrey, & Crowley, 2014; Richey et al., 2015]. Given the enhanced sensitivity of the EDI, it is possible that using it in conjunction with neuroimaging or genetic analyses may help identify important endophenotypes.

In sum, the EDI provides an efficient, valid, and sensitive method to measure emotion dysregulation in youth with ASD that may prove informative for screening and treatment monitoring as well as phenotypic characterization in biologically focused studies. It is noteworthy that the EDI items, particularly those for Reactivity, provided more information than related and commonly used longer measures, including the ABC-I and CBCL scales, even when utilizing the 7-item EDI short form. Further, the lack of gender, IQ, and verbal ability biases makes the EDI a rare tool that can be used across the full spectrum of functioning in ASD.

Acknowledgements

Subjects were recruited and data obtained in partnership with the Interactive Autism Network (IAN) Research Database at the Kennedy Krieger Institute, Baltimore, MD. IAN collaborating investigators and staff included: Paul H. Lipkin, MD, J. Kiely Law, MD, MPH, Alison R. Marvin, PhD. Data were also collected in partnership with the Autism and Developmental Disorders Inpatient Research Collaborative (ADDIRC) through use of Autism Inpatient Collection (AIC) data. The ADDIRC is made up of the co-investigators: Matthew Siegel, MD (PI) (Maine Medical Center Research Institute; Tufts University), Craig Erickson, MD (Cincinnati Children's Hospital; University of Cincinnati), Robin L. Gabriels, PsyD (Children's Hospital Colorado; University of Colorado), Desmond Kaplan, MD and Rajeesh Mahajan, MD (Sheppard Pratt Health System), Carla Mazefsky, PhD (Western Psychiatric Institute and Clinic; University of Pittsburgh), Eric M. Morrow, MD, PhD (Bradley Hospital; Brown University), Giulia Righi, PhD (Bradley

Hospital; Brown University), Susan L Santangelo, ScD (Maine Medical Center Research Institute; Tufts University), and Logan Wink, MD (Cincinnati Children's Hospital; University of Cincinnati). Collaborating investigators and staff: Jill Benevides, BS, Carol Beresford, MD, Carrie Best, MPH, Katie Bowen, LCSW, Briar Dechant, BS, Tom Flis, BCBA, LCPC, Holly Gastgeb, PhD, Angela Geer, BS, Louis Hagopian, PhD, Benjamin Handen, PhD, BCBA-D, Adam Klever, BS, Martin Lubetsky, MD, Kristen MacKenzie, BS, Zenoa Meservy, MD, John McGonigle, PhD, Kelly McGuire, MD, Faith McNeil, BS, Joshua Montrenes, BS, Tamara Palka, MD, Ernest Pedapati, MD, Kahsi A. Pedersen, PhD, Christine Peura, BA, Joseph Pierri, MD, Christie Rogers, MS, CCC-SLP, Brad Rossman, MA, Jennifer Ruberg, LISW, Elise Sannar, MD, Cathleen Small, PhD, Nicole Stuckey, MSN, RN, Barbara Tylanda, PhD, Brittany Troen, MA, RDMT, Mary Verdi, MA, Jessica Vezzoli, BS, and Deanna Williams, BA. We would like to extend a special thank you to IAN and AIC families for making this study possible. We would also like to thank Nate Dodds for his assistance in preparing the figures. This study was supported with funding from NICHD R01 HD079512 (CM), the Ritvo-Slifka Award for Innovation in Autism Research (CM), Simons Foundation Autism Research Initiative (SFARI #296318, M.S.) and the Nancy Lurie Marks Family Foundation. IAN is a partnership project of the Kennedy Krieger Institute and the Simons Foundation. IAN is also partially funded through a Patient-Centered Outcomes Research Institute (PCORI) Award for development of the National Patient-Centered Clinical Research Network, known as PCORnet

Conflicts of Interest

The authors declare no conflicts of interest.

References

- Achenbach, T.M., & Rescorla, L. (2001). Manual for the ASEBA school-age forms & profiles: An integrated system of multi-informant assessment. Burlington, VT: University of Vermont, Research Center for Children, Youth & Families.
- Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review*, 30, 217–237.
- Aman, M.G., Singh, N.N., Stewart, A.W., & Field, C.J. (1985a). The Aberrant Behavior Checklist: A behavior rating scale for the assessment of treatment effects. *American Journal of Mental Deficiency*, 89, 485–491.
- Aman, M.G., Singh, N.N., Stewart, A.W., & Field, C.J. (1985b). Psychometric characteristics of the aberrant Behavior Checklist. *American Journal of Mental Deficiency*, 89, 492–502.
- Ashburner, J., Ziviani, J., & Rodger, S. (2010). Surviving in the mainstream: Capacity of children with autism spectrum

- disorders to perform academically and regulate their emotions and behavior at school. *Research in Autism Spectrum Disorders*, 4, 18–27.
- Chandler, S., Howlin, P., Simonoff, E., O'Sullivan, T., Tseng, E., Kennedy, J., ... & Baird, G. (2016). Emotional and behavioural problems in young children with autism spectrum disorder. *Developmental Medicine & Child Neurology*, 58, 202–208.
- Choi, S.W. (2009). Firestar: Computerized adaptive testing (CAT) simulation program for polytomous IRT models. *Applied Psychological Measurement*, 33, 644–645.
- Choi, S.W., Reise, S.P., Pilkonis, P.A., Hays, R.D., & Cella, D. (2010). Efficiency of static and computer adaptive short forms compared to full-length measures of depressive symptoms. *Quality of Life Research*, 19, 125–136.
- Cicchetti, D., Ackerman, B., & Izard, C. (1995). Emotions and emotion regulation in developmental psychopathology. *Development and Psychopathology*, 7, 1–10.
- Cicchetti, D., & Toth, S.L. (1991). A developmental perspective on internalizing and externalizing disorders. In: Cicchetti, D. & Toth, S.L. (Eds.), *Rochester symposium on developmental psychopathology*, vol 16. internalizing and externalizing expressions of dysfunction (pp. 1–19) Hillsdale, NJ: Lawrence Erlbaum.
- 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.
- Coppin, G., & Sander, D. (2016). "Theoretical approaches to emotion and its measurement". In H. Meiselman (Ed.), *Emotion measurement* (pp. 3–30), Cambridge: Woodhead Publishing.
- Daniels, A.M., Rosenberg, R.E., Anderson, C., Law, J.K., Marvin, A.R., & Law, P.A. (2012). Verification of parent-report of child autism spectrum disorder diagnosis to a web-based autism registry. *Journal of Autism and Developmental Disorders*, 42, 257–265.
- Embretson, S.E., & Reise, S.P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Fernandez, K.C., Jazaieri, H., & Gross, J.J. (2016). Emotion regulation: A transdiagnostic perspective on a new RDoC domain. *Cognitive Therapy and Research*, 40, 426–440.
- Griffin, C., Lombardo, M.V., & Auyeung, B. (2016). Alexithymia in children with and without autism spectrum disorders. *Autism Research*, 9, 773–780.
- Gross, J.J., (Ed.). (2013). *Handbook of emotion regulation* (669 p) New York: Guilford.
- Norcross, J.C., & Wampold, B.E. (2010). What works for whom: Tailoring psychotherapy to the person. *Journal of Clinical Psychology*, 67, 1–6.
- Interactive Autism Network (IAN). (2017). Linking the autism community and researchers. Retrieved 8 September 2017, from <https://www.ianresearch.org/>.
- Irwin, D.E., Gross, H.E., Stucky, B.D., Thissen, D., DeWitt, E.M., Lai, J.S., ... DeWalt, D.A. (2012). Development of six PROMIS pediatrics proxy-report item banks. *Health and Quality of Life Outcomes*, 10, 22.
- Jahromi, L.B., Meek, S.E., & Ober-Reynolds, S. (2012). Emotion regulation in the context of frustration in children with high functioning autism and their typical peers. *Journal of Child Psychology and Psychiatry*, 53, 1250–1258.
- Janicke, D.M., Finney, J.W., & Riley, A.W. (2001). Children's health care use: A prospective investigation of factors related to care-seeking. *Medical Care*, 39, 990–1001.
- Kaat, A.J., Lecavalier, L., & Aman, M.G. (2014). Validity of the Aberrant Behavior Checklist in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 44, 1103–1116.
- Konstantareas, M.M., Stewart, K., (2006). Affect regulation and temperament in children with autism spectrum disorder. *Journal of autism and developmental disorders*, 36(2), 143–154.
- Larsen, R.J., & Diener, E. (1992). Promises and problems with the circumplex model of emotion. In Clark, M.S. (Ed.), *Review of Personality and Social Psychology*, No. 13. Emotion. Thousand Oaks, CA: Sage, pp. 25–59.
- Lerner, M.D., White, S.W., ... McPartland, J.C. (2012). Mechanisms of change in psychosocial interventions for autism spectrum disorders. *Dialogues in clinical neuroscience* 14(3), 307.
- Marvin, A.R., Law, P.A., Law, J.K., Arthur, R.M., Mortenson, E.M., Abbacchi, A.M., ... Constantino, J.N. (2014). Non-verbal children with ASD (NV-ASD): validating a registry and characterizing a population. *International Meeting for Autism Research*.
- Mayes, S.D., Kokotovich, C., Mathiowetz, C., Baweja, R., Calhoun, S.L., & Waxmonsky, J. (2017). Disruptive Mood Dysregulation Disorder Symptoms by Age in Autism, ADHD, and General Population Samples. *Journal of Mental Health Research in Intellectual Disabilities*, 10, 345–359. <https://doi.org/10.1080/19315864.2017.1338804>
- Mazefsky, C.A., Borue, X., Day, T.N., & Minshew, N.J. (2014). Emotion regulation patterns in adolescents with high-functioning autism spectrum disorder: Comparison to typically developing adolescents and associations with psychopathology. *Autism Research*, 7, 344–354.
- Mazefsky, C.A., Day, T.N., Siegel, M., White, S.W., Yu, L., & Pilkonis, P.A. (2016). Development of the emotion dysregulation inventory: A PROMIS® ing method for creating sensitive and unbiased questionnaires for autism spectrum disorder. *Journal of Autism and Developmental Disorders*, Advance Online Publication, DOI 10.1007/s10803-016-2907-1.
- Mazefsky, C.A., Herrington, J., Siegel, M., Scarpa, A., Maddox, B., Scahill, L., & White, S.W. (2013). The role of emotion regulation in autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 52, 679–688.
- Mikita, N., Hollocks, M.J., Papadopoulos, A.S., Aslani, A., Harrison, S., Leibenluft, E., ... Stringaris, A. (2015). Irritability in boys with autism spectrum disorders: An investigation of physiological reactivity. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 56, 1118–1126.
- Muthen, L.K., & Muthen, B.O. (2010) *Mplus User's Guide*, (4th ed.) Los Angeles, CA: Muthén & Muthén.
- Nuske, H.J., Hedley, D., Woollacott, A., Thomson, P., Macari, S., & Dissanayake, C. (2017). Developmental delays in emotion regulation strategies in preschoolers with autism. *Autism Research*, DOI: 10.1002/aur.1827

- Lee, H., Marvin, A.R., Watson, T., Piggot, J., Law, J.K., Law, P.A., Constantino, J.N., & Nelson, S.F. (2010). Accuracy of Phenotyping of Autistic Children Based on Internet Implemented Parent Report. *American Journal of Medical Genetics, Part B*, 153B, 1119–1126.
- Lord, C., Risi, S., Lambrecht, L., Cook, E.H., Jr., Leventhal, B.L., DiLavore, P.C., et al. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30, 205–223.
- Patel, S., Day, T.N., Jones, N., & Mazefsky, C.A. (2016). Association between anger rumination and autism symptom severity, depression symptoms, aggression, and general dysregulation in adolescents with autism spectrum disorder. *Autism: International Journal of Research and Practice*, 21, 181–189.
- Pitskel, N.B., Bolling, D.Z., Kaiser, M.D., Pelphrey, K.A., & Crowley, M.J. (2014). Neural systems for cognitive reappraisal in children and adolescents with autism spectrum disorder. *Developmental Cognitive Neuroscience*, 10, 117–128.
- Pouw, L.B., Rieffe, C., Stockmann, L., & Gadow, K.D. (2013). The link between emotion regulation, social functioning, and depression in boys with ASD. *Research in Autism Spectrum Disorders*, 7, 549–556.
- Prizant, B.M., Wetherby, A.M., Rubin, E., & Laurent, A.C. (2003). The SCERTS Model: A transactional, family-centered approach to enhancing communication and socioemotional abilities of Children with Autism Spectrum Disorder. *Infants & Young Children*, 16, 296–316.
- Ravindran, N., & Myers, B.J. (2012). Cultural influence on perceptions of health, illness, and disability: A review and focus on autism. *Journal of Child and Family Studies*, 21, 311–319.
- Richey, J.A., Damiano, C.R., Sabatino, A., Rittenberg, A., Petty, C., Bizzell, J., ... Davidson, R.J. (2015). Neural mechanisms of emotion regulation in autism spectrum disorder. *Journal of autism and developmental disorders*, 45, 3409–3423.
- Rieffe, C., Oosterveld, P., Terwogt, M.M., Mootz, S., Van Leeuwen, E., & Stockmann, L. (2011). Emotion regulation and internalizing symptoms in children with autism spectrum disorders. *Autism*, 15, 655–670.
- Roid, G.H., Miller, L.J., Pomplun, M., & Koch, C. (2013). *Leiter international performance scale*, (3rd ed.): Leiter-3. Wood Dale, IL: Stoelting Co.
- Rothbart, M.K. (2006). Temperament, development, and personality. *Current Directions in Psychological Science*, 16, 207–212.
- Rothbart, M.K., & Sheese, B.E. (2007). Temperament and emotion regulation. In J. Gross (Ed.), *Handbook of emotion regulation* (pp. 331–350). New York, NY: Guilford Press.
- Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 29, 1161–1178.
- Rutter, M., Bailey, A., & Lord, C. (2003). *Social communication questionnaire (SCQ)*. Los Angeles, CA: Western Psychological Services.
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph*, 17, 1969;
- Samson, A.C., Hardan, A.Y., Podell, R.W., Phillips, J.M., & Gross, J.J. (2015). Emotion regulation in children and adolescents with autism spectrum disorder. *Autism Research*, 8, 9–18.
- Samson, A.C., Hardan, A.Y., Lee, I.A., Phillips, J.M., & Gross, J.J. (2015). Maladaptive behavior in autism spectrum disorder: the role of emotion experience and emotion regulation. *Journal of Autism and Developmental Disorders*, 45, 3424–3432.
- Samson, A.C., Huber, O., & Gross, J.J. (2012). Emotion regulation in Asperger's syndrome and high-functioning autism. *Emotion*, 12, 659.
- Samson, A.C., Phillips, J.M., Parker, K.J., Shah, S., Gross, J.J., & Hardan, A.Y. (2014). Emotion dysregulation and the core features of autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 44, 1766–1772.
- Schäfer, J.Ö., Naumann, E., Holmes, E.A., et al. (2017). Emotion Regulation Strategies in Depressive and Anxiety Symptoms in Youth: A Meta-Analytic Review. *Journal of Youth Adolescence*, 46, 261.
- Siegel, M., Smith, K.A., Mazefsky, C., Gabriels, R.L., Erickson, C., Kaplan, D., ... Santangelo, S.L. for the Autism and Developmental Disorders Inpatient Research Collaborative (ADDIRC). (2015). The autism inpatient collection: Methods and preliminary sample description. *Molecular Autism*, 6, 61–71.
- Shields, A.M., & Cicchetti, D. (1997). Emotion regulation among school-age children: The development and validation of a new criterion Q-sort scale. *Developmental Psychology*, 33, 906–916.
- Teresi, J.A., Ocepek-Welikson, K., Kleinman, M., Eimicke, J.P., Crane, P.K., Jones, R.N., ... Reise, S.P. (2009). Analysis of differential item functioning in the depression item bank from the Patient Reported Outcome Measurement Information System (PROMIS): An item response theory approach. *Psychology Science Quarterly*, 51, 148.
- Thissen, D., Steinberg, L., & Wainer, H. 1993. Detection of differential item functioning using the parameters of item response models. In: Holland, P.W., Wainer, H. (Eds.), *Differential item functioning* (pp. 67–113). Hillsdale, NJ: Lawrence Erlbaum.
- Ting, V., & Weiss, J.A. (2017). Emotion regulation and parent co-regulation in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 47, 680–689.
- Weiss, J.A., Riosa, P.B., Mazefsky, C.A., & Beaumont, R. (2017). Emotion Regulation in Autism Spectrum Disorder. In Essau, C.A., LeBlanc, S.S., & Ollendick, T.H. (Eds.), *Emotion regulation and psychopathology in children and adolescents*, Oxford: Oxford University Press, pp. 235–258.
- Weiss, J.A., Thomson, K., & Chan, L. (2014). A Systematic Literature Review of Emotion Regulation Measurement in Individuals With Autism Spectrum Disorder. *Autism Res*, 7, 629–648. doi:10.1002/aur.1426
- White, S.W., Lerner, M.D., McLeod, B.D., Wood, J.J., Ginsburg, G.S., Kerns, C., ... Compton, S. (2015). Anxiety in youth with and without autism spectrum disorder: examination of factorial equivalence. *Behavior Therapy*, 46, 40–53.
- White, S.W., Mazefsky, C.A., Dichter, G.S., Chiu, P.H., Richey, J.A., & Ollendick, T.H. (2014). Social-cognitive,

- physiological, and neural mechanisms underlying emotion regulation impairments: Understanding anxiety in autism spectrum disorder. *International Journal of Developmental Neuroscience*, 39, 22–36.
- White, S.W., Elias, R., Salinas, C., Capriola, N., Conner, C.M., Asselin, S., ...Getzel, E. (2016). Students with Autism Spectrum Disorder in College: Results from a Preliminary Mixed Methods Needs Analysis. *Research in Developmental Disabilities*. doi: 10.1016/j.ridd.2016.05.010.,
- Zantinge, G., van Rijn, S., Stockmann, L., & Swaab, H. (2017). Physiological Arousal and Emotion Regulation Strategies in Young Children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 47, 2648–2657.
- Zelkowitz, R.L., & Cole, D.A. (2016). Measures of emotion reactivity and emotion regulation: Convergent and discriminant validity. *Personality and Individual Differences*, 102, 123–132.
- Zumbo, B.D. (1999). A handbook on the theory and methods of differential item functioning (dif): Logistic regression modeling as a unitary framework for binary and likert-type (ordinal) item scores. Ottawa, ON: Directorate of Human Resources Research and Education, Department of National Defense.