Article

The Flexible Regulation of Emotional Expression Scale for Youth (FREE-Y): Adaptation and Validation Across a Varied Sample of Children and Adolescents

Assessment
2023, Vol. 30(4) 1265–1284

DOI: 10.1177/10731911221090465
journals.sagepub.com/home/asm
SAGE

Ann-Christin Haag[1], Christine B. Cha[1], Jennie G. Noll[2,3], Dylan G. Gee[4], Chad E. Shenk[2,5], Hannah M. C. Schreier[2], Christine M. Heim[2,6], Idan Shalev[2], Emma J. Rose[2], Alana Jorgensen[1], and George A. Bonanno[1]

Abstract

Flexible self-regulation has been shown to be an adaptive ability. This study adapted and validated the adult *Flexible Regulation of Emotional Expression* (FREE) Scale for use with youth (FREE-Y) in community and maltreatment samples. The FREE-Y measures the ability to flexibly enhance and suppress emotion expression across an array of hypothetical social scenarios. Participants (N = 654, 8–19 years) were included from three studies. Confirmatory factor analysis (CFA) confirmed a theoretically appropriate higher order factor structure. Using multiple-group CFAs, measurement invariance was achieved across maltreatment status, age, and gender. Reliabilities were adequate and construct validity was demonstrated through associations with measures of emotion regulation, psychopathology, IQ, and executive functioning. Group comparisons indicated lower Suppression and Flexibility scores for maltreated versus comparison participants. Findings suggest that the FREE-Y is a valid measure of expressive regulation ability in youth that can be applied across a range of populations.

Keywords

emotion regulation, flexibility, children and adolescents, youth, psychometric evaluation, invariance, maltreatment

Mature emotion regulation ability has been characterized by the flexibility to deploy specific strategies in a manner that effectively matches changing situational demands (Skinner & Zimmer-Gembeck, 2007). These skills play a crucial role in positive adjustment across the life span (reviewed in Berking & Wupperman, 2012; Zeman et al., 2006). Yet, although flexible emotion regulation has received increasing attention in the adult literature (Aldao et al., 2015; Benson et al., 2019; Bonanno & Burton, 2013), in particular due to its role in adjustment after adverse events (Bonanno et al., 2004; Westphal et al., 2010), it is still relatively under-researched in youth. This gap is especially problematic as it pertains to the flexible regulation of emotional expression, which represents an important milestone in adolescent development (Zeman et al., 2006). Accordingly, to address this deficit in the current study, we adapted the *Flexible Regulation of Emotional Expression* (FREE) Scale developed for adults for use with preadolescent and adolescent samples.

Expressive Flexibility

The ability to flexibly enhance or suppress emotional expression is known as *expressive flexibility (EF)* (Westphal et al., 2010). Work in this domain developed in response to traditional perspectives on emotion regulation that tended to view regulatory strategies, including the expression and

[1]Columbia University Teachers College, New York, USA
[2]The Pennsylvania State University, University Park, USA
[3]Cincinnati Children's Hospital Medical Center, University of Cincinnati, OH, USA
[4]Yale University, New Haven, CT, USA
[5]The Pennsylvania State University College of Medicine, Hershey, USA
[6]Charité—Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Germany

Corresponding Author:
Ann-Christin Haag, Department of Counseling and Clinical Psychology, Teachers College, Columbia University, 525 West 120th Street, Box 218, New York, NY 10025, USA.
Email: ah3784@tc.columbia.edu

suppression of emotion, in relatively static terms, as either uniformly adaptive or maladaptive (reviewed in Aldao et al., 2010). As researchers began to explore these propensities across situational variations, it became increasingly evident that consequences of emotional expression and suppression tend to vary by situation (Bonanno et al., 2007; Coifman & Bonanno, 2010; Kalokerinos et al., 2017).

Bonanno and colleagues (2004) proposed that "whether one expresses or suppresses emotional expression is not as important for adjustment as is the ability to flexibly express or suppress emotional expression as demanded by the situational context" (p. 482). These researchers developed an experimental paradigm to measure that ability, the EF task, in which participants enhanced emotional expression, suppressed emotional expression, and behaved naturally on different trials (Bonanno et al., 2004). More recently, a compatible self-report measure, the FREE Scale (Burton & Bonanno, 2016), was developed and validated against the experimental paradigm. The scale was based on a scenario approach and, mirroring the EF task, showed a higher order factor structure consisting of two second-order factors—expressive enhancement ability and suppression ability—representing four first-order factors grouped by regulation and emotional valence: enhance positive, enhance negative, suppress positive, and suppress negative. In addition to the enhancement and suppression scores, an overall EF score can be created.

Research using these measures in adult populations has linked EF with multiple facets of emotional and psychological well-being, such as lower scores of depression, anxiety, and higher life satisfaction (Burton & Bonanno, 2016; S. Chen et al., 2018). Furthermore, deficits in EF have been found in individuals with psychopathology (Gupta & Bonanno, 2011; Rodin et al., 2017). Although the concept of EF has been well established in adults, only very little research has been done so far in youth.

Expressive Flexibility in Youth

Learning to regulate emotions has been identified as one of the most important tasks starting in early childhood. Preschool age (3–5 years) is an important period for developing skills in theory of mind and perspective taking (Carlson et al., 2013; Zelazo, 2015). At the very young age of 3 years, children start to be able to regulate their expressive behaviors, which has been shown in the very commonly used paradigm of the "disappointing present" (Saarni, 1984). The ability to regulate expressivity improves through childhood and is related to learning to differentiate between emotion and expression (Kromm et al., 2015). As emotional expression emerges in the context of social interaction (Holodynski & Friedlmeier, 2006), children have to learn about display rules. The understanding and use of display rules increase over elementary school age (Zeman et al., 2006). At the ages 7–12 years, children have been shown to adapt their expression of emotions in response to the expectations of others in different interpersonal contexts (Zeman & Garber, 1996). Furthermore, by school age, children show evidence of metacognition and flexible use of metacognitive strategies (Davis et al., 2010). Although the precision with which children tailor their choices of emotion regulation strategies according to different situations is thought to increase, by the age of 7 to 11 years, children already show a sophisticated repertoire of emotion regulation strategies that they can apply flexibly (Parsafar et al., 2019).

As emotion regulation skills evolve through childhood, adolescence, and into adulthood, they become increasingly differentiated and autonomous, eventually peaking in the mature ability to flexibly match regulation strategies to meet situational demands (Skinner & Zimmer-Gembeck, 2007). To date, a limited number of studies has conceptualized emotion flexibility in youth in various ways, such as, for example, switching between strategies to regulate emotions or switching between emotional states, or the display of different affective behaviors (Hollenstein et al., 2004; Parsafar et al., 2019; Van der Giessen et al., 2015). In each of these conceptualizations, greater levels of emotion flexibility have been linked to more positive mental health outcomes. However, relatively little research has been conducted with EF.

The lack of research on EF in youth is even more concerning as the appropriate regulation of emotion expression is widely viewed as an important milestone in middle childhood and adolescence (Zeman et al., 2006). Heightened emotional reactivity coupled with the awareness of the interpersonal consequences for a particular display of emotion and with elevated sensitivity to the evaluations of others (Stifter & Augustine, 2019; Zeman et al., 2006) renders emotion expression a crucial developmental skill, for example, for establishing positive social relationships (Wang et al., 2020).

Recently, the experimental EF task was adopted to measure youth's expressive enhancement, suppression, and flexibility (Wang & Hawk, 2019). However, a direct adaptation of the FREE Scale from adults to youth is still lacking. Wang and Hawk (2020) did develop a related self-report instrument assessing EF in youth revealing a two-factor structure of enhancement and suppression (Wang & Hawk, 2020). However, this measure raised several methodological concerns. First, the items were not directly derived, that is, adapted, from the adult FREE Scale. Second, this study was normed in Chinese youth only, suggesting the possibility that some items might not be very well suited for a Western sample (e.g., "When a classmate does something funny in class, I can refrain from laughing to help maintain order"). Finally, they did not evaluate their EF measure in a sample of youth exposed to

extreme adversity, which, as we discuss below, has been shown to form a crucial component of the construct's validity in adult samples.

Flexibility and Adversity

A number of studies have linked EF in adults with improved psychological adjustment in the aftermath of adverse events (Bonanno et al., 2004; Westphal et al., 2010). Of particular relevance, one study examined EF in a sample of adult survivors of childhood maltreatment. Using the experimental EF task (Bonanno et al., 2004), the authors found that a history of childhood maltreatment was associated with reduced levels of EF (Pițur & Miu, 2020).

There are two rationales for why EF warrants special attention in youth who experienced maltreatment. First, children who experienced abuse or neglect exhibit lower levels of emotion regulation than non-maltreated children (Kim-Spoon et al., 2013). Research has also shown that children exposed to maltreatment have greater difficulties with the recognition, expression, and understanding of emotions (reviewed in Assed et al., 2020). Second, there is empirical evidence to suggest that mature emotion regulation skills, thus potentially EF, can be beneficial in youths' posttraumatic adjustment after maltreatment. For example, children who experienced childhood maltreatment but had achieved more mature levels of emotion regulation had lower levels of internalizing symptoms over time (Kim & Cicchetti, 2010). Complementarily, female adolescent survivors of sexual abuse who had greater emotion regulation difficulties were found to have more severe posttraumatic stress and depressive symptoms (Chang et al., 2018). To date, however, no study has investigated the role of EF in the development of psychopathology in youth after maltreatment. Together, this evidence suggests that it is imperative to adapt and validate the adult FREE self-report scale for use in youth and to include a sample of survivors of childhood maltreatment as part of that validation process.

The Current Study

Addressing these gaps, the present study adapted the items from the original FREE Scale for adults to capture corresponding scenarios appropriate for youth. The newly adapted *Flexible Regulation of Emotional Expression Scale for Youth* (FREE-Y) was evaluated in several U.S. samples, including subsamples of participants with recent records of investigations for childhood maltreatment, and participants with a history of suicidal ideation, and a comparison group of youth from community samples. Using these data, we aimed to (a) interrogate the FREE-Y and examine its factor structure and reliability in the full sample; (b) test measurement invariance across maltreatment status, age, and gender (in a reduced sample excluding youth with a history of suicidal ideation); (c) compare mean FREE-Y scores (Enhancement, Suppression, and Flexibility) across these three groups; and (d) evaluate the scale's validity by investigating its associations to other theoretically and clinically relevant variables.

As the adult version, the FREE-Y Scale was designed to produce a multifactor structure consisting of the regulatory abilities of expressive enhancement and suppression in both positively and negatively valenced emotions. Because the FREE-Y Scale was designed using these predefined theoretical factors, we conducted confirmatory factor analyses (CFAs), rather than exploratory analyses, to compare model fit indices across competing nested models that differ in their regulation ability type and valence-type factor structures. We expect adequate fit of the higher order model presented in Burton and Bonanno (2016) to our data, as well as adequate internal consistencies. We further expect the multigroup test to reveal measurement invariance across maltreatment status, age, and gender. Based on the existing literature presented in the introduction (Assed et al., 2020; Kim & Cicchetti, 2010; Kim-Spoon et al., 2013), we anticipate group differences between participants investigated for child maltreatment versus comparison participants, in terms of the maltreatment group reporting lower means in EF scores. Previous studies have shown some gender differences in EF abilities using experimental measures: one study of EF in youth (Wang & Hawk, 2019) and a mixed pattern has been observed in adults. However, many studies have not examined gender differences in flexibility and no study has examined gender differences using the adult FREE Scale. Thus, because the FREE-Y is a novel measure, we propose no formal hypotheses regarding gender differences but will conduct exploratory analyses. Similarly, we also chose to examine age differences in an exploratory manner. Research on EF across disparate age groups in adults has not observed significant age differences (e.g., Emery & Hess, 2011; Wang & Hawk, 2019). Again, however, because the FREE-Y is a novel measure, we decided to include no specific hypotheses on age differences on this measure in a youth sample. With regard to construct validity, we chose an array of constructs based on their theoretical relevance and based on what the two existing studies examining self-report scales to assess EF have reported (Burton & Bonanno, 2016; Wang & Hawk, 2020) for the sake of comparability of their results with the FREE-Y. Thereby, we chose widely used measures of emotion regulation as they also assess facets of self-regulation, just like EF. We expect positive associations with measures of emotion regulation and negative associations with instruments assessing difficulties in emotion regulation. Furthermore, based on previous literature in adults showing that EF is an adaptive ability (Bonanno & Burton, 2013; S. Chen et al., 2018; Gupta & Bonanno, 2011; Rodin et al., 2017) and based on the results of the existing studies examining

Table 1. Sample Characteristics of the Full Combined Sample (N = 654) and the Reduced Sample (n = 571).

Characteristics		Full combined sample		Reduced sample[a]	
		N/*M*	%/*SD*	*N*/*M*	%/*SD*
Age		13.01	2.98	12.30	2.47
Sex	Female	341	53.0	288	51.4
	Male	289	45.0	267	47.7
	Other	13	2.0	5	0.9
Race	White	381	59.2	357	63.7
	Black/African American	102	15.8	86	15.4
	Asian	54	8.4	27	4.8
	Multiracial	62	9.6	57	10.2
	Other	39	6.1	29	5.2
	Unknown	6	0.9	4	0.7
Ethnicity	Hispanic	106	16.5	83	14.8
	Non-Hispanic	535	83.5	477	85.2
Household family income	<$25,000	221	40.8	218	42.2
	$25,000–$49,999	131	24.2	125	24.2
	$50,000–$74,999	74	13.7	72	13.9
	$75,000–$100,000	35	6.5	34	6.6
	>$100,000	81	14.9	68	13.2
Type of investigated maltreatment (Study 1, *n* = 373)	Sexual abuse	57	15.3	57	15.3
	Physical abuse	168	45.0	168	45.0
	Neglect	148	39.7	148	39.7

[a]The reduced sample, that is, excluding those participants who reported a history of suicidal ideation (*n* = 83), was used for the measurement invariance testing across maltreatment status and for the group comparisons of participants who have experienced maltreatment versus comparison participants.

self-report scales of EF (Burton & Bonanno, 2016; Wang & Hawk, 2020), we included a set of measures of common psychopathologies (depression, anxiety, posttraumatic stress disorder [PTSD]) to investigate construct validity and expect negative correlations with EF. Finally, as previous research in adults suggests that EF processes require cognitive resources, thus limiting memory capacities (Bonanno et al., 2004), we wanted to elucidate this relationship further in youth and expect EF to be positively associated with a standardized measure of IQ and negatively with difficulties in executive functioning.

Method

Participants and Procedure

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. The present study combines data from across three sites that administered the FREE-Y and numerous measures of emotion regulation, psychopathology (e.g., anxiety, depression), IQ, and executive functioning among youth up until September 2020. All study procedures were approved by the Institutional Review Boards at Columbia University's Teachers College, Yale University, and Pennsylvania State University, and all participants provided informed age-appropriate assent (youth < 18 years) and consent (caregiver) prior to enrollment in the study. The consenting/assenting process was carried out in a private area, by trained staff. It was stated that subjects are free to discontinue participation at any time, without penalty. In addition to Human Subjects training, all staff involved with the recruitment or collection of data from subjects received specific training in the administration and importance of informed consent.

Overall, 654 participants aged 8 to 19 years were included, referred to as "youth" in the present study. Table 1 displays sociodemographic data of the full combined study sample.

Although the primary purpose and recruitment strategies varied across sites (detailed below), it was necessary to combine samples from the three studies (a) to examine both clinical and comparison samples, considering differences in EF after adverse events and its potential buffering role in adjustment; (b) to include a wide age range spanning childhood to adolescence; and (c) to attain sufficient power for the analyses. Although Studies 1 and 2 include both participants considered as either clinical (i.e., having a history of child maltreatment or suicidal ideation) or serving as comparison groups, Study 3 was an unselected community

sample. The subsample of youth who reported a history of suicidal ideation was not analyzed separately as its sample size was limited ($n = 83$).

Study 1. The first sample featured the first 439 children enrolled into a large longitudinal cohort study aimed at examining the impact of child maltreatment (i.e., sexual abuse, physical abuse, and child neglect) on subsequent adverse health consequences (see Schreier et al., 2021, for full details). As part of this cohort study, children aged 8 to 13 ($M = 11.42$, $SD = 1.44$, 47.9% female) who experienced an investigation for child maltreatment within the previous year were recruited via Statewide Child Welfare Information System (CWIS) records. A demographically matched comparison group of children were also enrolled. Children with developmental disabilities (reported by caregivers and/or confirmed through CWIS), who are not able to read and understand English (as per caregiver report), whose CWIS involvement was more than 12 months ago, and whose caregivers are not able to read and understand English or refuse to participate are not eligible. Children in the comparison group additionally are not eligible if CWIS screening results in any records of prior involvement with Child Welfare agencies. Although recruitment for this cohort study is ongoing with a target enrollment of 900 children, this initial subset of 439 was utilized for the purposes of the psychometric evaluation of the FREE-Y.

Study 2. The second sample featured 172 adolescents (12–19 years, $M = 17.05$, $SD = 1.97$, 67.3% female) recruited from the community, drawn from a study examining cognitive risk factors for suicidal ideation. Participants were screened for suicidal ideation before enrolling them in the study and 48.3% ($n = 83$) of the sample included in the present study has a history of suicidal ideation. Study exclusion criteria were as follows: the presence of any factor impairing the adolescent's ability to effectively participate in the study, including non-fluency in English, the presence of gross cognitive impairment, agitated/violent behavior, or high/imminent risk of suicide. Participants are recruited via online advertisements and using flyers in the New York City community.

Study 3. The third sample featured 43 older children and adolescents (8–17 years, $M = 13.11$, $SD = 2.62$, 48.6% female) recruited from the community, drawn from a study examining transdiagnostic processes related to stress and anxiety (e.g., emotional learning and regulation). Participants were recruited using flyers and online advertisements in the New Haven area. Exclusion criteria included (a) cognitive impairment (Full-Scale Intelligence Quotient <80); (b) history of head injury or concussion; (c) history of chronic medical illness or neurological disorder; (d) lifetime history of psychotic disorders, autism spectrum disorder, bipolar disorder, conduct disorder, non-alcohol or non-tobacco substance use disorder, current alcohol or tobacco use disorder, current primary diagnosis of attention-deficit/hyperactivity disorder, or major depressive disorder; (e) acute suicidal ideation; (f) current use of psychotropic medication; (g) color blindness; (h) visual impairment that cannot be corrected; and (i) hearing impairment. Contraindications for n magnetic resonance imaging (MRI) scan (e.g., braces, metal implants) and left-handedness were also exclusionary given that a separate component of the study involved an MRI scan.

Measures

Expressive Flexibility. The *Flexible Regulation of Emotion Expression Scale for Youth* (FREE-Y) was developed for the present investigation and administered across all three sites. The FREE-Y Scale was designed to assess children and adolescents' perceived ability to modulate (i.e., enhance and suppress) their displayed emotion across different hypothetical scenarios. Scenarios from the adult version of the FREE Scale (Burton & Bonanno, 2016) were translated to real-life scenarios that youth can identify with and have most probably experienced. Items from the adult version of the FREE Scale (Burton & Bonanno, 2016) were adapted according to their face validity by researchers (AJ and CBC), one of whom (CBC) has worked extensively with children and adolescents in educational, clinical, and research contexts. FREE items were adapted in two ways: by removing scenarios that some adolescents may not yet have experienced (e.g., replace "date" with "fun weekend"; replace work-related experiences with school), and by adjusting language to be more age appropriate (e.g., replace "conceal" with "hide"). Overall, nine of the 16 FREE items were adapted to children and adolescents, while seven remained unchanged. For participants of Study 1, who range between 8 and 13 years of age, instructions and items of the FREE-Y were read to them from a script by trained and probed research assistants to ensure they understood everything correctly. The items are rated on a 6-point scale (1 = *unable*, 6 = *very able*). Four expressive abilities are assessed by clusters of four items each: *enhancing positive emotion*, *enhancing negative emotion*, *suppressing positive emotion*, and *suppressing negative emotion*. An *EF score* was obtained by subtracting the absolute value of the difference between the enhancement and suppression scores from a sum flexibility score, which was calculated by adding up enhancement and suppression scores (Westphal et al., 2010).

Alternative Measures of Emotion Regulation. Three well-established measures of emotion regulation, not focusing on emotion expression abilities, were administered in the present investigation. The *Emotion Regulation Checklist* (ERC; Shields & Cicchetti, 1998) and the *Emotion*

Regulation Index for Children and Adolescents (ERICA; MacDermott et al., 2010) were administered to Study 1, and the *Modified Difficulties in Emotion Regulation Scale* (M-DERS; Bardeen et al., 2016) was administered to Studies 2 and 3. Specifically, the ERC is a 24-item caregiver-report questionnaire of youth emotion regulation and emotion reactivity. Items were rated on a 4-point Likert-type scale (1 = *never*, 4 = *almost always*) assessing the frequency of emotional lability, expression, and regulation for the index child. Internal consistency for the ERC was excellent in the present study (α = .91). In addition to the ERC, the ERICA is a 16-item, self-report measure assessing emotion regulation during childhood and adolescence along three facets (i.e., emotional control, emotional self-awareness, and situational responsiveness). Participants rated how much each statement is true about them on a 5-point scale (1 = *strongly disagree*, 5 = *strongly agree*). Internal consistency for the total score used in the present study was adequate (α = .78). Finally, the M-DERS, which is based on the *Difficulties in Emotion Regulation Scale* (Gratz & Roemer, 2004), is a 29-item self-report measure assessing five dimensions of emotion regulation (i.e., Identification, Non-acceptance, Impulse, Goals, Strategies). Participants rated to what extent each phrase applies to them on a 5-point scale (1 = *almost never*, 5 = *almost always*). Internal consistency was excellent for the total score in the present study (α = 96).

Anxiety. The *Screen for Child Anxiety Related Disorders* (SCARED; Birmaher et al., 1997), a 41-item validated and reliable questionnaire that assesses anxiety disorder symptoms in children and adolescents, was administered across all three study sites using the child self-report version. Participants rated how true a statement is for them on a 3-point Likert-type scale (0 = *not true or hardly ever true*, 2 = *very true or often true*). A cutoff of ≥25 may indicate the presence of an anxiety disorder. This well-established measure was used in all three samples included in the present study and results in a total anxiety symptom severity score (α = .94), as well as five factors common to pediatric anxiety disorders: panic (α = .88), generalized anxiety (α = .87), separation anxiety (α = .79), social anxiety (α = .84), and school avoidance (α = .71).

Depressive Symptoms. Depressive symptoms were captured using *z*-transformed scores from two self-report measures: the *Children's Depression Inventory–2* (CDI-2; Kovacs, 2015), administered to Studies 1 and 3, and the *Quick Inventory of Depressive Symptomatology–Self-Report* (QIDS-SR; Rush et al., 2003), administered to Study 2. The CDI-2 contains 28 items assessing symptoms of pediatric depression in children and adolescents. Items assess the cognitive, affective, and behavioral dimensions of pediatric depression and are ranked on a 3-point Likert-type scale (0 = *low*, 2 = *high*). A cutoff of ≥20 may indicate clinically relevant levels of depression (Matthey & Petrovski, 2002). The overall internal consistency of the CDI-2 was good (α = .88). In addition, the QIDS-SR contains 16 items assessing recent symptoms of a major depressive episode by the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM-IV*; American Psychiatric Association [APA], 2000) on a 4-point Likert-type scale (0–3). Total scores ≥11 indicate moderate to severe depression (Bernstein et al., 2010). The internal consistency of the QIDS-SR was good (α = .82). CDI-2 and QIDS-SR scores were *z*-transformed and merged to obtain a depression score for the full combined sample of the present study.

Posttraumatic Stress. Symptoms of PTSD were captured using the *UCLA PTSD Reaction Index* (UCLA PTSD-RI; Pynoos, 2015). In this fully structured clinical interview with youth, participants rate the frequency of 31 PTSD symptoms over the past month on a 5-point Likert-type-scale (0 = *none of the time* to 4 = *most of the time*). A PTSD diagnosis was made based on *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM-5*; APA, 2013) criteria. The UCLA PTSD-RI was used in Study 1 and yielded an excellent internal consistency in the present study (α = .94).

IQ and Executive Functioning. Study 1 specifically featured the *Woodcock–Johnson IV Test of Cognitive Abilities* (WJ-IV). The WJ-IV is a standardized measure of overall and domain-specific aspects of intelligence (Schrank et al., 2014), and was administered by trained study staff to youth of Study 1. The overall assessment of intellectual ability, that is, IQ, was used in the present study. Participants of Study 1 also completed the *Behavior Rating Inventory of Executive Functioning–2* (BRIEF-2; Gioia et al., 2015) to assess pediatric executive functioning rated by the caregiver. Subscales measure domain-specific aspects of executive function, such as inhibition, shifting, emotion control, planning, and organizing, that can be used to generate indices of emotion regulation (α =.93), behavioral regulation (α = .92), and cognitive regulation (α = .96), as well as a global regulation index (α = .98). Caregivers rate the frequency of 63 executive functioning difficulties of their child on a 3-point Likert-type-scale (0 = *never* to 2 = *often*).

Statistical Analysis

Data were analyzed using the statistical software R version 3.6.2 (R Core Team, 2021) for Mac. CFA was applied using the R *lavaan* (0.5-20) package (Rosseel, 2012) and multiple-group CFA was applied using the R *semTools* (0.5-4.0) package (Jorgensen et al., 2021). CFA was used to examine the factor structure of the FREE-Y Scale within the full combined sample of the present study. Thereby, goodness of fit for four models was examined. First, a single-factor model

was tested, where all items loaded onto a single-regulation-type factor. Second, we tested an emotion-based two-factor model, with eight items loading on either a factor Enhance or Suppress. The third model tested included four factors, each comprising four items: enhance positive emotion, enhance negative emotion, suppress positive emotion, and suppress negative emotion. Finally, a fourth higher order model was tested, with four first-order factors (the same as described in the four-factor model) and two second-order factors, Enhance and Suppress. All CFAs were estimated using the robust maximum likelihood estimation procedure with robust standard errors and a scaled test statistic, due to non-normality of scores of several items (Rhemtulla et al., 2012). Using Little's test of missing completey at random (MCAR), missing data in the FREE-Y items can be assumed to be MCAR, $\chi^2(64) = 64.03$, $p = .475$. Therefore, missing data (1.6%) were replaced using full information maximum likelihood estimation. Model fit for the different CFAs was evaluated with the model χ^2, the comparative fit index (CFI), the robust root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). Although a non-significant χ^2 value ($p > .05$) theoretically implies that the tested model is consistent with the data, it has been discussed that this is rarely achieved as it is dependent on sample size (Cheung & Rensvold, 2002). Simulation studies state that CFI values close to or above 0.95, RMSEA less than 0.06, and SRMR values less than 0.08 indicate acceptable model fit (Hu & Bentler, 1999; Iacobucci, 2010; Maydeu-Olivares et al., 2018; Schermelleh-Engel et al., 2003).

The properties of the FREE-Y Scale were investigated across three different groups of participants: (a) Sample 1 participants investigated for maltreatment as identified via child welfare records versus a comparison group from all three samples (maltreatment status), (b) participants younger than 13 years versus 13 years and older (age), and (c) girls and boys (gender binary). To maintain a more homogeneous non-clinical comparison group, participants who reported having a history of suicidal ideation in Study 2 ($n = 83$) were excluded from the comparison group for the measurement invariance tests (and subsequent group comparisons). Regarding the age cutoff, 12 or 13 years are often referred to as the transition to adolescence (e.g., Casey, 2015; A. O. Cohen et al., 2016), and we ultimately chose 13 years as it aligns with the mean age of youth in our full combined sample.

Tests of measurement invariance were conducted with this reduced sample to investigate whether the same construct has been measured across groups of maltreatment, age, and gender. First, two baseline CFA models were fit for each group comparison. Following current recommendations (F. F. Chen et al., 2005; Rudnev et al., 2018), six models were examined to assess measurement invariance in a second-order factor model. Multiple-group CFA using the robust maximum likelihood estimator was computed for the following sequential models: *Configural Invariance*—The number of factors and the pattern of fixed and free factor loadings were constrained to be the same across groups. This configural model was then used to compare against the next more restrictive model in the sequence (Model 1). *Metric Invariance*—Factor loadings of first-order (Model 2a), and first- and second-order factors (Model 2b) were constrained to be equal across groups. *Scalar Invariance*—Item intercepts (Model 3a), and item and first-order intercepts (Model 3b) and first- and second-order factor loadings were constrained to be equal across groups. *Uniqueness*—Unique variances of first-order factors (Model 4) and intercepts, and first- and second-order factor loadings were constrained equally. Testing equal error variances of the items has been regarded as exceedingly stringent (Byrne, 2010) and therefore was not included in the present study. Changes in model fit resulting of the more and more stringent equality constraints were investigated by comparing the models sequentially. Comparisons of nested models were computed via Satorra–Bentler scaled χ^2-difference tests (Satorra & Bentler, 2001) and changes in CFI and RMSEA. As the χ^2-criterion has been considered too stringent as it is dependent on sample size (e.g., F. F. Chen, 2007; Cheung & Rensvold, 2002), generally accepted guidelines regarding changes in CFI and RMSEA were also applied in the present article. Accordingly, measurement invariance was considered established if the S-Bχ^2-difference test resulted in a p value $>.05$ or if a difference of $|\Delta CFI| \geq .010$ in the CFI was supplemented by a change of $|\Delta RMSEA| \geqslant .015$ in RMSEA (F. F. Chen, 2007; Cheung & Rensvold, 2002; Putnick & Bornstein, 2016). Both the total sample size used in the CFA with the full sample and the sample sizes per group in measurement invariance tests are considered adequate as per recommendations (Barrett, 2007; F. F. Chen, 2007; Cheung & Rensvold, 2002).

To examine differences in FREE-Y Scale scores across maltreatment status, age, and gender, two omnibus multivariate analyses of variance (MANOVAs) and subsequent univariate analyses of variance (ANOVAs) were conducted, that is, one model for maltreatment versus comparison using the reduced sample to have a more homogeneous comparison group (as described above) and one model for the binary variables of age and gender using the full sample. Results of the univariate ANOVAs were adjusted for multiple comparisons by means of rotation testing (Langsrud, 2005) using the R *ffmanova* (1.1.0) package (Langsrud & Mevik, 2019). Relations between the FREE-Y Scale and other validity measures were investigated using bivariate Pearson correlations. The study samples used for computing the correlations vary, as different sets of measures were administered to the participants of the three studies included in the present work. Guidelines to interpret effect sizes are as follows: partial $\eta^2 = .01$ or $r = .10$ indicates small, partial $\eta^2 = .06$ or $r = .30$ indicates medium, and partial $\eta^2 = .14$ or $r = .50$ indicates large effects (J. Cohen, 1988).

Results

Sample Descriptives

The full combined sample included 654 participants between the ages of 8.19 and 19.00 years (M = 13.01 years, SD = 2.98). As described above, a reduced sample was used in analyses where the participants who experienced maltreatment were analyzed against a comparison group. Characteristics of both the full combined and the reduced sample are displayed in Table 1. In addition, descriptive statistics and group comparisons across Studies 1 to 3 are presented in Table 2. While participants across Studies 1 to 3 differed in age, no significant differences were found for depression or anxiety (Table 2).

CFA

The fit of four alternative structural models of increased complexity was investigated in the full sample via CFA, presented in Table 3. Although both the single- and the two-factor models did not adequately fit the data, results of the four-factor model indicated that all indices of the measurement model were in an acceptable range. Furthermore, comparisons indicated significantly better fit for the four-factor model than for the two-factor model (Model C vs. B; Table 3). The higher order factor model evidenced similarly good fit to the data as the four-factor model. In practice, however, previous research has emphasized the importance of ability categories, rather than valence of the regulated emotion. For example, expressive enhancement and suppression abilities predicted later psychological adjustment, independent of valence of the emotion (Bonanno et al., 2004; Westphal et al., 2010). The primacy of ability over valence type is consistent with the higher order factor model. Thus, in all subsequent analyses, we focus on the higher order factor model and the second-order factors.

The higher order CFA model with the unconstrained factor loadings and intercepts is shown in Figure 1. All FREE-Y Scale's item loadings on their respective factors were good, except for Item 14 with a loading $<.40$.

Item Statistics and Reliability

Table 4 contains the FREE-Y Scale items and their item statistics in the full combined sample. The item means were of medium to large size. Almost all item-total correlations were good ($r_{it} > .40$), except those for Items 7 and 14. Internal consistencies were good for the eight-item composite Enhance ($\alpha = .77$, $M = 32.08$, $SD = 8.05$, skewness = −0.22, kurtosis = −0.30) and Suppress scales ($\alpha = .74$, $M = 28.86$, $SD = 8.45$, skewness = 0.18, kurtosis = −0.42), acceptable for the Enhance—positive ($\alpha = .73$, $M = 17.50$, $SD = 4.52$, skewness = −0.59, kurtosis = −0.17) and the Suppress—positive ($\alpha = .70$, $M = 15.30$, $SD = 5.02$, skewness = −0.03, kurtosis = −0.79) subscales, but were comparatively lower for the Enhance—negative ($\alpha = .65$, $M = 14.59$, $SD = 4.78$, skewness = −0.04, kurtosis = −0.66) and the Suppress—negative ($\alpha = .60$, $M = 13.56$, $SD = 4.85$, skewness = 0.13, kurtosis = −0.54) subscales.

Measurement Invariance

Multigroup CFA was performed to investigate whether the higher order model was invariant across maltreatment status, age, and gender. Results are displayed in Table 5.

Maltreatment Status Invariance. For the comparison of maltreated versus comparison participants (using the reduced sample, excluding the participants who were screened for suicidal ideation), the higher order factor model had acceptable fit when fitted separately in both groups. Configural invariance was supported by an acceptably fitting baseline model. This configural model was then used to compare against the more restrictive measurement invariance models, that is, metric invariance. Metric invariance was tested constraining factor loadings of first-order (Model 2a) and first- and second-order factors (Model 2b) following recommendations described above (F. F. Chen, 2007; F. F. Chen et al., 2005). Both models of metric invariances fit the data well. Changes in CFI and RMSEA, when the metric invariance models were compared with previous models in the sequence, were within acceptable values (ΔCFI = .000 and –.001, ΔRMSEA = −.001 and .000). Metric invariance across maltreatment status indicates that the items used to estimate the factor loadings have the same meaning for youth who experienced maltreatment and comparison participants. The next restrictive models, that is, the models of first-order scalar (Model 3a) and first- and second-order scalar (Model 3b) invariance, also fit the data well. Scalar invariance was demonstrated as changes in CFI and RMSEA were again within acceptable values (ΔCFI = −.006 and .001, ΔRMSEA = −.001 and .000), even though the S-Bχ^2-difference test was significant for Model 3a. This indicates that first- and second-order factor loadings, item intercepts, and first-order intercepts are invariant across maltreatment status. Finally, the most restrictive model, that is, invariance of unique variances (Model 4), was rejected according to the S-Bχ^2-difference rule, but not according to ΔCFI and ΔRMSEA rules (ΔCFI = −.004, ΔRMSEA = .001). In summary, measurement invariance was established, suggesting that the FREE-Y measures the same construct across maltreatment status and that average item score comparisons are valid between youth who experienced maltreatment and comparison participants.

Age Invariance. For the analyses of age invariance (<13 years vs. 13 years and older) using the full combined sample, the configural models showed good fit to the data.

Table 2. Descriptive Statistics and Study Group Comparisons Across Outcomes Assessed in Studies 1 to 3.

	Study 1 ($n = 439$)				Study 2 ($n = 172$)				Study 3 ($n = 43$)				
Variable	*M*	*SD*	Minimum–maximum	≥Cutoff[a]	*M*	*SD*	Minimum–maximum	≥Cutoff[a]	*M*	*SD*	Minimum–maximum	≥Cutoff[a]	*F*- or *t*-statistics
Age	11.42	1.44	8.25–14.00	—	17.05	1.97	12.00–19.00	—	13.11	2.62	8.19–17.74	—	$F(2, 650) = 684.35$, adjusted $p < .001$
Emotion Regulation Difficulties (M-DERS)	—	—	—	—	70.05	22.80	29.00–128.00	—	55.54	23.37	29.00–129.00	—	$t = 3.62$, $p < .001$
Emotion Regulation (ERC)	74.78	11.32	37.00–96.00	—	—	—	—	—	—	—	—	—	—
Emotion Regulation (ERICA)	55.79	9.29	25.00–80.00	—	—	—	—	—	—	—	—	—	—
PTSD symptoms	21.12	16.53	0.00–71.00	26.7	—	—	—	—	—	—	—	—	—
Anxiety total score	27.67	16.60	0.00–74.00	53.4	27.98	15.23	1.00–67.00	56.4	23.60	15.09	0.00–62.00	39.5	$F(2, 628) = 1.34$, adjusted $p = .390$
Depression[b]	11.13	8.72	0.00–48.00	15.7	8.14	4.88	0.00–22.00	28.8	10.21	9.90	0.00–50.81	11.4	$F(2, 634) = 0.24$, adjusted $p = .795$
IQ	98.14	17.24	40.00–147.00	—	—	—	—	—	—	—	—	—	—
Global regulation index	104.91	27.01	60.00–179.00	—	—	—	—	—	—	—	—	—	—

Note. M-DERS = Modified Difficulties in Emotion Regulation Scale; ERC = Emotion Regulation Checklist; ERICA = Emotion Regulation Index for Children and Adolescents; PTSD = posttraumatic stress disorder; MANOVA = multivariate analysis of variance.

[a]Percentage (%) equal to or above the cutoff indicating levels of clinical relevance of anxiety and depression. For PTSD, the % of PTSD diagnosis according to the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM-5*; APA, 2013) is presented.

[b]Depression was assessed via the *Children's Depression Inventory* in Studies 1 and 3, and via the *Quick Inventory of Depressive Symptomatology* in Study 2. For the comparison of depression across Studies 1 to 3, *z* scores were used. MANOVA: $F(6, 1238) = 109.03$, $p < .001$. Descriptive statistics for anxiety in the full combined sample were as follows: $M = 26.68$, $SD = 16.17$, minimum–maximum = 0.00–74.00, percentage ≥cutoff = 50.9.

Table 3. Results of the Confirmatory Factor Analyses of the Different Models.

Models	χ^{2a}	*df*	CFI	RMSEA	90% CI	SRMR	vs.	Δdf	$\Delta\chi^2$
A. Single-factor	535.96	104	.740	.090	[.083, .098]	.071			
B. Two-factor	300.74	103	.881	.061	[.053, .069]	.051	A	1	171.39***
C. Four-factor	188.34	98	.946	.042	[.033, .051]	.043	B	5	95.87***
D. Higher order	189.17	99	.946	.042	[.033, .051]	.043	C	1	0.88

Note. $N = 645$ (full combined sample). Robust fit indices presented. CFI = comparative fit index; RMSEA = root mean square error of approximation; 90% CI = 90% confidence interval for RMSEA; SRMR = standardized root mean square residual; Δdf = difference in degrees of freedom.
[a]All *ps* $< .001$.
***$p < .001$.

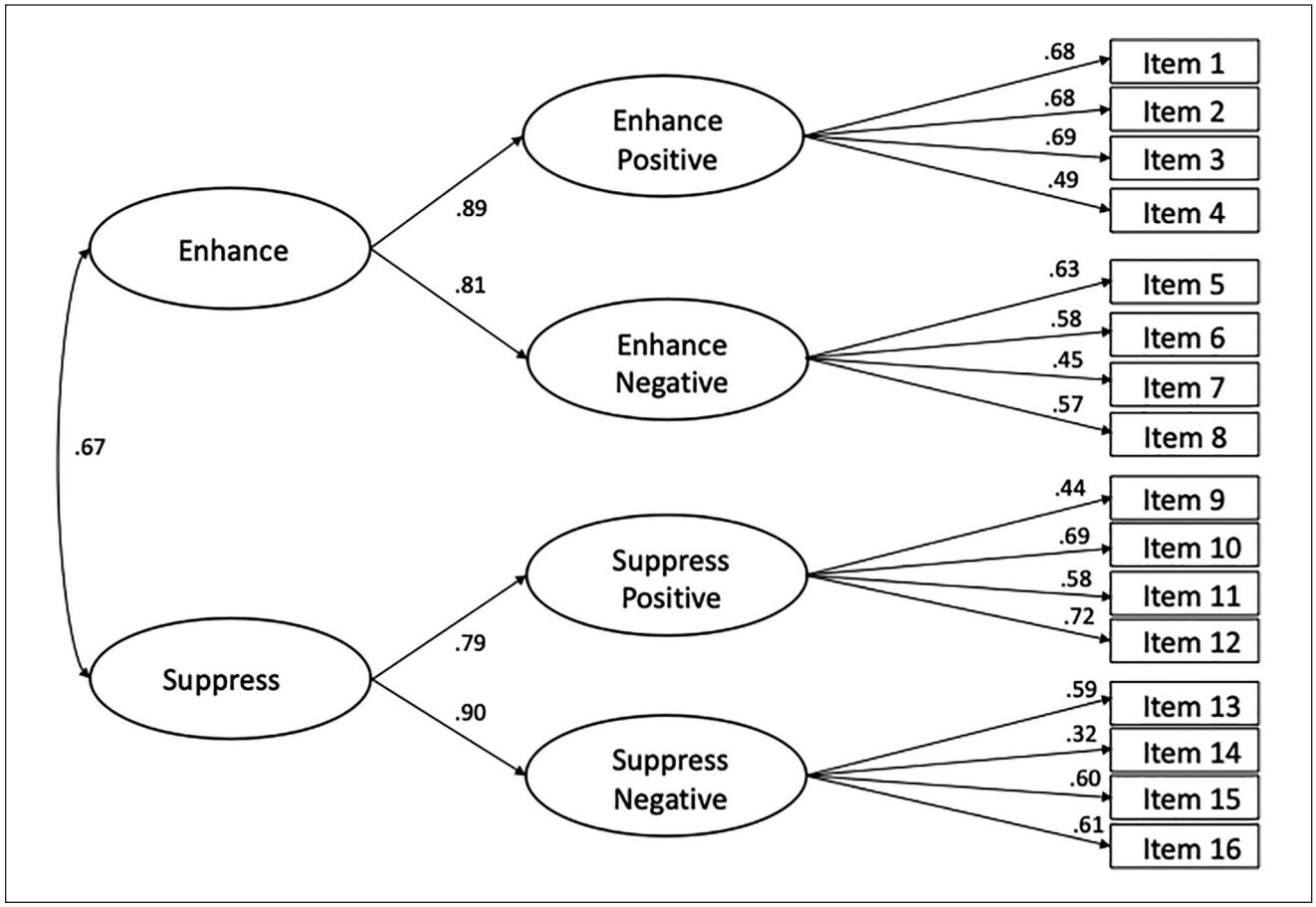

Figure 1. Higher-Order Factor Model of the FREE-Y.
Note. Standardized factor loadings presented from confirmatory factor analysis. All factor loadings are substantial and significant at $p < .01$. $N = 654$. FREE-Y = Flexible Regulation of Emotional Expression Scale for Youth.

Applying the S-Bχ^2-differences rule, first-order scalar invariance of the item intercepts (Model 3a) was rejected across the age groups. However, all model comparisons showed invariance across age, considering changes in CFI and RMSEA.

Gender Invariance. Invariance across gender (girls vs. boys) was established in the full combined sample. The baseline model fitted the data well, indicating configural invariance. The subsequent nested models demonstrated metric and scalar invariance as well as invariance of unique variances applying both the S-Bχ^2-differences and ΔCFI and ΔRMSEA rules.

In summary, despite contrasting results using the S-Bχ^2-difference tests for three model evaluations, the FREE-Y can be considered invariant across youth who experienced maltreatment and comparison participants, across younger and older youth, as well as across girls and boys at the most

Table 4. Item Descriptives of the 16 Items of the FREE-Y Scale.

Instructions and items (1 = *unable*, 6 = *very able*)	*M*	*SD*	r_{it}
Enhance positive			
"*The following scenarios involve POSITIVE emotion. For each scenario, indicate how well you would be able to be even MORE EXPRESSIVE than usual of how*":			
1 A friend wins an award for a sport that doesn't interest you.	4.57	1.49	.62
2 A friend gets a good grade and wants to talk about it.	4.36	1.52	.61
3 A friend is talking about the fun weekend they had.	4.43	1.54	.61
4 You receive a gift from a family member, but it's a shirt you dislike.	4.13	1.57	.49
Enhance negative			
"*The following scenarios involve NEGATIVE emotion. For each scenario, indicate how well you would be able to be even MORE EXPRESSIVE than usual of how*":			
5 Your friend is telling you about what a terrible day they had.	4.27	1.61	.59
6 Your teacher is complaining about another class they teach that you do not take and know little about.	3.16	1.70	.52
7 A friend is talking about a bad fight they had with another friend that you secretly think is a good thing.	3.60	1.74	.39
8 You're attending the funeral of someone you don't know.	3.56	1.79	.53
Suppress positive			
"*The following scenarios involve POSITIVE emotion. For each scenario, indicate how well you would be able to HIDE how you were feeling*":			
9 While having lunch with a friend who has just recently failed a class, you find out that you earned honor roll.	4.42	1.57	.42
10 You are in class and you see an accidentally funny misspelling the teacher writes on the board.	3.69	1.82	.61
11 You're a guest at a serious religious ceremony and the person sitting next to you just whispered a funny joke.	3.79	1.71	.52
12 During a meeting with a teacher, the teacher's phone unexpectedly begins to play a funny and embarrassing ringtone.	3.40	1.84	.65
Suppress negative			
"*The following scenarios involve NEGATIVE emotion. For each scenario, indicate how well you would be able to HIDE how you were feeling*":			
13 You are at a social event and the person you're talking to frequently spits while they speak.	3.34	1.79	.53
14 You hear that a close family member passed away just before you have to give an important class presentation.	3.10	1.84	.29
15 You are having lunch with a new classmate on the first day of school, and a stranger spills their drink on you.	3.31	1.84	.52
16 After you have a very irritating and stressful day, a sometimes annoying neighbor stops by to say hello.	3.81	1.72	.56

Note. N = 642–645 (full combined sample). r_{it} = corrected item-total correlation per each of the two FREE-Y second-order factors; FREE-Y = Flexible Regulation of Emotional Expression Scale for Youth.

restrictive levels of measurement invariance applying ΔCFI and ΔRMSEA rules.

Group Differences

Mean differences for the FREE-Y Scale scores, Enhance, Suppress, and Flexibility, were examined across groups. The reduced sample was used when examining differences across maltreatment status, whereas the full combined sample was used for comparisons across gender and age. For the comparison of maltreatment versus comparison groups—Omnibus MANOVA, $F(3, 557) = 3.01$, $p = .030$—although no significant differences were found for the Enhance score, $F(1, 561) = 2.32$, adjusted $p = .131$, partial $\eta^2 = .004$, there were significant differences between participants of the maltreatment versus the comparison group in the Suppress, $F(1, 560) = 9.28$, adjusted $p = .014$, partial $\eta^2 = .02$, and Flexibility scores, $F(1, 560) = 6.32$, adjusted $p = .036$, partial $\eta^2 = .01$. For both scores, participants investigated for maltreatment had significantly lower scores (Suppress: $M = 27.73$, $SD = 8.59$; Flexibility: $M = 51.80$, $SD = 16.09$) than the comparison participants (Suppress: $M = 30.01$, $SD = 8.11$; Flexibility: $M = 55.28$, $SD = 14.67$). For gender, Omnibus multivariate analysis of covariance (MANCOVA), $F(3, 615) = 6.49$, $p < .001$, significant differences were found in the Enhance scores, $F(1, 618) = 13.18$, adjusted $p = .002$, partial $\eta^2 = .02$, with girls reporting greater Enhancement abilities ($M = 33.14$, $SD = 7.44$) than boys ($M = 30.72$, $SD = 8.51$). No significant differences were found between female and male participants in the Suppress,

Table 5. Goodness-of-Fit Indices and Model Comparisons for Measurement Invariance Models Across Maltreatment, Age, and Gender.

Invariance	Model	S-Bχ^{2a}	*df*	CFI	RMSEA	90% CI	SRMR	vs.	Δdf	ΔS-Bχ^2	ΔCFI	ΔRMSEA
Maltreatment[b]	Maltreatment (n = 373)	158.01	99	.932	.044	[.031, .057]	.050					
	Comparison (n = 198)	147.11	99	.933	.054	[.034, .072]	.058					
	1. Configural	305.25	198	.933	.048	[.037, .058]	.053					
	2a. First-order metric	318.29	210	.933	.047	[.036, .057]	.057	1	12	12.03	.000	–.001
	2b. First- and second-order metric	322.51	212	.931	.047	[.036, .057]	.059	2a	2	4.50	–.001	.000
	3a. First-order scalar	345.19	224	.925	.047	[.037, .057]	.060	2b	12	23.42*	–.006	.001
	3b. First- and second-order scalar	349.23	226	.924	.048	[.037, .057]	.061	3a	2	4.22	–.001	.000
	4. Unique variances	360.43	230	.920	.048	[.039, .058]	.063	3b	4	11.53*	–.004	.001
Age	<13 years (n = 392)	149.44	99	.948	.040	[.026, .052]	.048					
	13 years and older (n = 261)	142.74	99	.945	.046	[.028, .062]	.053					
	1. Configural	292.12	198	.946	.042	[.032, .052]	.050					
	2a. First-order metric	303.09	210	.948	.041	[.030, .050]	.052	1	12	9.20	.001	–.002
	2b. First- and second-order metric	307.13	212	.946	.041	[.030, .051]	.054	2a	2	4.11	–.001	.000
	3a. First-order scalar	338.55	224	.936	.043	[.034, .053]	.056	2b	12	34.11***	–.010	–.003
	3b. First- and second-order scalar	239.47	226	.937	.043	[.033, .052]	.056	3a	2	0.63	.001	.000
	4. Unique variances	341.39	230	.938	.042	[.033, .051]	.056	3b	4	2.02	.001	–.001
Gender	Female (n = 341)	160.58	99	.918	.048	[.034, .061]	.053					
	Male (n= 289)	112.14	99	.985	.024	[.000, .043]	.044					
	1. Configural	273.87	198	.953	.039	[.027, .050]	.049					
	2a. First-order metric	289.88	210	.951	.039	[.027, .049]	.053	1	12	15.89	–.002	.000
	2b. First- and second-order metric	290.55	212	.952	.038	[.026, .048]	.053	2a	2	0.27	.001	–.001
	3a. First-order scalar	310.69	224	.947	.039	[.028, .049]	.055	2b	12	20.82	–.005	.001
	3b. First- and second-order scalar	316.11	226	.945	.039	[.028, .049]	.055	3a	2	5.84	–.002	.001
	4. Unique variances	318.07	230	.946	.039	[.028, .048]	.056	3b	4	2.08	.001	–.001

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; 90% CI = 90% confidence interval for RMSEA; SRMR = standardized root mean square residual; Δdf = difference in degrees of freedom between nested models; ΔS-Bχ^2 = difference in ΔS-Bχ^2 between nested models.

[a]For measurement invariance testing across maltreatment status, the reduced sample was used, excluding participants who reported a history of suicidal ideation (n = 571). Robust fit indices presented. S-Bχ^2 = Satorra–Bentler scaled statistic for nested model comparisons.

[b]All $ps \leq .001$ or $p < .01$.

*$p < .05$. ***$p < .001$.

$F(1, 617) = 0.10$, adjusted $p = .750$, and Flexibility scores, $F(1, 617) = 3.47$, adjusted $p = .098$. No significant differences were found for the age group comparisons, Omnibus MANCOVA, $F(3, 615) = 2.11$, $p = .098$.

Association of the FREE-Y Scale With Other Measures of Emotion Regulation, Psychopathology, IQ, and Executive Functioning

Bivariate correlations of the FREE-Y Scale's first- and second-order factors with various validity measures of emotion regulation, psychopathology, and executive functioning (according to their availability in the three included studies) are presented in Table 6.

Expressive Enhancement. Self-reported ability to enhance their emotional expression showed significant small to medium correlations with emotion regulation assessed via the ERICA (sample from Study 1). Although the direction of association between enhancement and self-reported difficulties in emotion regulation (M-DERS; Studies 2 and 3) and frequency of emotion regulation rated by the caregiver (ERC; Study 1) were as hypothesized, these associations were not statistically significant. Expressive enhancement was significantly negatively correlated with depression. No significant associations were found with PTSD, anxiety, problematic executive functioning, and IQ (all from Study 1).

Expressive Suppression. Significant positive correlations in a modest range were found for the suppression of emotional expression with self-rated emotion regulation assessed by the ERICA, whereas no such correlations were found for caregiver-rated emotion regulation assessed via the ERC (both from Study 1). The M-DERS subscale Impulse was significantly negatively associated with Suppression (Studies 2 and 3). In addition, Suppression was significantly negatively correlated with depression and school avoidance in the full combined sample, but not with PTSD (Study 1). A significant modest positive association was found with IQ (Study 1). Regarding difficulties in executive functioning, a significant negative association was revealed with the Emotion Regulation Index, whereas no significant association was revealed with the Global Regulation Index, and the Cognitive and the Behavior Regulation Index (Study 1).

Expressive Flexibility. A significant medium-size positive correlation was found between Flexibility and self-rated emotion regulation along the ERICA (Study 1). In addition, Flexibility correlated significantly and negatively with both total emotion regulation difficulties and the Impulse subscale (M-DERS; Studies 2 and 3). Flexibility showed small but significant negative correlations with depression, social anxiety, and school avoidance (full sample), and was significantly positively associated with IQ (Study 1). No significant associations were found between Flexibility and PTSD, executive functioning, and caregiver-rated emotion regulation along the ERC (all from Study 1).

Post hoc exploratory analyses on variances were conducted to investigate a potential moderating role of maltreatment status in the association of EF (Enhance, Suppress, Flexibility) and the depression and those anxiety scales (Social Anxiety, School Avoidance) that have been shown to be significantly associated with EF. While no significant interaction terms were found for all FREE-Y scores in the prediction of Depression and School Avoidance (all ps $> .05$), maltreatment was a (marginally) significant moderator of the association between each of the FREE-Y scores Suppression and Flexibility and Social Anxiety: Suppression, $F(1, 545) = 4.40$, $p = .036$; Flexibility, $F(1, 545) = 3.04$, $p = .081$.

Discussion

The FREE-Y showed good psychometric properties, specifically an appropriate higher order factor structure; measurement invariance across maltreatment status, age, and gender; adequate reliabilities; and construct validity evidenced in correlations with theoretically relevant markers of emotion regulation, psychopathology, IQ, and executive functioning. Together, these findings suggest that the newly adapted FREE-Y is a reliable and valid measure of EF in youth.

Consistent with the findings for the adult FREE (Burton & Bonanno, 2016), CFAs revealed similarly good fit for the four-factor and the higher order factor structure of the FREE-Y. Item loadings on the respective factors in the hierarchal factor model were generally good for the FREE-Y, with the exception of one item. This item also had the lowest item-total correlation, suggesting that this item may be less clearly related to the other items, as well as to the Suppress negative factor, and potentially less adequate for youth than for adults. However, because the factor loading and item-total correlation were close to the acceptable threshold, we retained this item in the FREE-Y to replicate the adult FREE Scale as much as possible.

Configural, metric, and scalar measurement invariance, as well as invariance of the unique variances of first-order factors have been established according to ΔCFI and ΔRMSEA rules, but not entirely based on the S-Bχ^2-difference rule. Leaning on the commonly used CFI and RMSEA criteria, these findings suggest that the FREE-Y captures the same construct across maltreatment status, age, and gender. These findings further suggest that the units and origins of the scale are stable, which permits

Table 6. Zero-Order Correlations of First- and Second-Order Factors of the FREE-Y Scale With Relevant Measures.

	Second-order factors			First-order factors			
Measure	Enhance	Suppress	Flexibility	Enhance positive	Enhance negative	Suppress positive	Suppress negative
Emotion regulation							
Emotion regulation difficulties (M-DERS; Studies 2 and 3, $n = 215$)	−.108	−.092	−.152*	−.186**	−.004	−.151*	−.002
Identification	−.098	.047	−.083	−.115	−.052	−.059	.138
Non-acceptance	−.075	−.056	−.096	−.160*	.024	−.095	.002
Impulse	−.082	−.191**	−.206**	−.171*	.024	−.207***	−.112
Goals	−.075	−.113	−.115	−.117	−.015	−.116	−.072
Strategies	−.103	−.052	−.110	−.180*	−.003	−.127	.040
Emotion regulation (ERC; Study 1, $n = 439$)	.042	.084	.063	.012	.060	.047	.097*
Emotion regulation (ERICA; Study 1, $n = 439$)	.237***	.263***	.305***	.263***	.155**	.212***	.241***
Psychopathology							
PTSD symptoms (Study 1, $n = 439$)	.011	−.037	−.055	.015	.005	−.009	−.056
Anxiety total score	.009	−.066	−.055	.020	−.003	−.060	−.053
Panic disorder	.031	−.059	−.033	.024	.030	−.060	−.040
Generalized anxiety	.048	−.021	−.009	.063	.022	−.029	−.006
Separation anxiety	.015	−.048	−.035	.033	−.007	−.017	−.065
Social anxiety	−.074	−.060	−.094*	−.050	−.077	−.064	−.038
School avoidance	−.019	−.114**	−.089*	−.021	−.012	−.083*	−.112**
Depression[a]	−.106**	−.137**	−.168***	−.139***	−.046	−.145***	−.089*
IQ and executive functioning (Study 1, $n = 439$)							
IQ	.085	.128**	.110*	.051	.096*	.075	.146**
Global regulation index	−.040	−.068	−.048	−.010	−.058	−.075	−.041
Behavior regulation index	−.050	−.092	−.084	−.022	−.063	−.082	−.076
Emotion regulation index	−.041	−.103*	−.075	.015	−.084	−.093	−.084
Cognitive regulation index	−.023	−.031	−.014	−.013	−.027	−.050	−.002

Note. FREE-Y Scale = Flexible Regulation of Emotional Expression Scale for Youth; M-DERS = Modified Difficulties in Emotion Regulation Scale; Pearson Correlation; ERC = Emotion Regulation Checklist; ERICA = Emotion Regulation Index for Children and Adolescents; PTSD = posttraumatic stress disorder.

[a]Z scores derived from Children's Depression Inventory (Studies 1 and 3) and Quick Inventory of Depressive Symptomatology (Study 2). Unless otherwise noted in parenthesis, the full combined sample was used ($n = 654$).

*$p < .05$. **$p < .01$. ***$p < .001$.

analyses of relations between constructs and comparisons of group means.

Comparing across groups, we found that participants exposed to maltreatment had lower Suppression and overall Flexibility abilities than comparison participants. Although high levels of adversity and/or psychopathology are generally associated with reduced flexibility in adults (reviewed in Coifman & Summers, 2019), particularly in the context of child maltreatment (Pițur & Miu, 2020), the associations between the specific ability scales and adjustment have varied across contexts, populations, and types of outcomes. For example, veterans with PTSD were found to have equal suppression ability as veterans without PTSD, but less enhancement ability (Rodin et al., 2017), whereas bereaved adults suffering from complicated grief were less able to both enhance and suppress emotional expression compared with asymptomatic or married adults (Gupta & Bonanno, 2011).

The disrupting effects of childhood maltreatment on emotion regulation development are well documented (Kim & Cicchetti, 2010; Noll, 2021). There are several possible explanations for the lowered suppression abilities among maltreated youth in the present study. First, maltreated youth show increased impulsivity. It has been hypothesized that neurocognitive adaptations, such as reduced stress reactivity, diminished cognitive and/or executive functioning capacities, in response to adverse rearing environments would exacerbate impulsivity (Lovallo, 2013; Oshri et al., 2018; Thibodeau et al., 2015). Second, maltreatment has been associated with diminished cognitive resources, which are necessary for expressive suppression (Bonanno et al., 2004; Richards & Gross, 2006). One reason for the diminished cognitive resources is that early life stress induced by maltreatment, as well as adverse rearing environments, has been linked to alterations in brain development (Morris et al., 2007; Noll, 2021). Another potential reason is that repeatedly threatening environments can induce high levels of distress, arousal, and vigilance that would compete for cognitive resources with suppression (Kim & Cicchetti, 2010). Both potential explanations are supported by the associations observed in the present study linking suppression inversely to IQ and the Impulse subscale of the M-DERS.

Unexpectedly, despite differences in FREE-Y scores across maltreatment status, we found no significant associations with PTSD symptomatology. The adult FREE showed negative associations with PTSD (Rodin et al., 2017) and, more generally, emotion regulation problems have been associated with greater levels of PTSD symptoms in adolescents (reviewed in Villalta et al., 2018). One possible explanation for the null association in the current study, however, may be that the levels of PTSD symptoms in our sample were relatively low, with $M = 21.25$ ($SD = 16.62$, median = 18.00, minimum = 0.00, maximum = 71.00) and a possible range of scores between 0 and 80. In addition, the associations with measures of emotion regulation indicate that the FREE-Y is theoretically consistent with our expectations in terms of significant negative correlations with emotion regulation difficulties (M-DERS) and significant positive associations with youth-rated emotion regulation (ERICA). However, we found no significant associations between caregiver-rated emotion regulation (ERC) and the FREE-Y. This fits well into patterns of large effect discrepancies between parent and youth reports on subscales of emotion regulation that have been shown to get more pronounced as youth's age increases (Hourigan et al., 2011). The discrepancies were also investigated across age in the present study, but the same patterns of correlations between the FREE-Y scores and the ERICA or the ERC were found for younger and older youth of Study 1 (8–13 years).

The FREE-Y Suppress and Flexibility scores showed inverse associations with the anxiety subscales for Social Anxiety and School Avoidance. This finding might be explained by previously observed links between reduced suppression abilities in youth and increased peer rejection (Wang & Hawk, 2020). Emotional expression is a mean of engaging in social relationships, a skill crucial in adolescence. If the youth who are less able to suppress their emotions have more difficulties with peers and in consequence are not as well accepted among peers as others, this might be exacerbating social and school anxieties. These mechanisms might, in turn, be reinforced by the fact that, especially early in adolescence, emotion regulation has been shown to be less successful in social compared to nonsocial situationsand if adolescents reported greater rejection sensitivity (Silvers et al., 2012).

Regarding the associations of the FREE-Y with other relevant measures used in the present study, it can be summarized that associations are mostly of low to moderate size. This is well in line with findings of both the adult FREE version (Burton & Bonanno, 2016) and the other existing scale measuring EF in youth (Wang & Hawk, 2020). In more detail, investigating associations with commonly used measures of emotion regulation, correlation coefficients do not exceed .30 in both the adult FREE (Burton & Bonanno, 2016) and the FREE-Y. Commonly used emotion regulation questionnaires, such as the ERC (Shields & Cicchetti, 1998) and the M-DERS (Bardeen et al., 2016) used in the present study, assess frequency of behavior rather than ability. It has been reasoned that regulation abilities, as measured by the FREE-Y, and regulation frequency are distinct constructs (Burton & Bonanno, 2016), thus explaining the low correlations in the present study. In addition, the magnitudes of the FREE-Y's associations with depression also match very well onto the results reported about the adult FREE (Burton & Bonanno, 2016).

To situate our findings in the broader literature on EF and psychopathology, we conducted post hoc exploratory

analyses to investigate a potential moderating role of maltreatment status in the association of EF and depression and anxiety. The existing literature has shown that EF is most clearly linked to psychopathology in the context of significant adverse events (Bonanno, 2005). Although no significant interaction terms were found in the prediction of Depression and School Avoidance in our study, maltreatment was a (marginally) significant moderator of the association between two facets of the FREE-Y, Suppression and Flexibility, and Social Anxiety. These findings provide further support for the validity of the FREE-Y.

Consistent with previous findings about EF in youth (Wang & Hawk, 2019) and adults (Westphal et al., 2010), we found no significant differences for the FREE-Y across age groups. However, we did find significant effects for gender, with girls reporting greater Enhancement abilities than boys. This is in contrast to findings of the only two existing studies investigating EF in youth that reported no gender differences in Enhancement (Wang & Hawk, 2019, 2020). From a broader perspective, though, it aligns well with existing findings regarding gender differences in emotion expression and regulation. In general, women have been found to use more emotion regulation strategies than men (reviewed in Nolen-Hoeksema, 2012). Furthermore, studies have shown that women express greater levels of emotions than men. Theories of gender socialization postulate that this is due to the adoption of gender-related display rules for emotion expression (reviewed in Chaplin & Aldao, 2013). Finally, research suggests that girls exhibit more agreeable personality traits (Klimstra et al., 2009), which supports our finding of girls being more able to enhance their emotional expression in the social scenarios presented by the FREE-Y.

Although this is the first study to assess the psychometric properties of a self-report measure of EF in youth with different experiences, several limitations of our study are worth noting. First, no information was available about potential involvement with child welfare services for Studies 2 and 3. However, to represent a homogeneous comparison group to participants who experienced maltreatment, we merged the comparison participants of Study 1 (who were screened for the absence of a child welfare record), the community-recruited participants of Study 2, explicitly excluding those screened for a history of suicidal ideation, and the community participants of Study 3. Furthermore, the results of both measurement invariance and group difference tests across youth investigated for maltreatment and comparison participants suggest adequate group allocation. A second limitation is that we did not have a large enough sample of youth screened for a history of suicidal ideation to analyze measurement invariance across those participants as a separate group. Findings may thus have limited generalizability to suicidal youth. Third, unlike the study of the adult version of the FREE, we were unable to validate the FREE-Y against experimental data from the EF task in this study. Fourth, we caution against the use of the FREE-Y in its current form with younger children, as this would likely require further adaptation of items and response scales.

Our findings of divergent associations of regulatory abilities with IQ and executive functioning indicate potential differences in the cognitive underpinnings of suppression and enhancement. Hence, the future study of cognitive mechanisms of flexibility is emphasized. Further research is also warranted investigating the predictive validity of the FREE-Y with clinical outcomes and its moderating effects among youth experiencing different types of adverse events. The use of the FREE-Y will enable clinicians to identify and monitor both resources and deficits in EF in youth, either representing possible protective or risk factors. The latter could be used as targets for interventions aiming at improving EF to support psychological well-being in general and importantly to foster psychological adjustment in the aftermath of adverse events. Presumably being a mechanism of resilience, fostering flexible self-regulation would entail fostering resilience (Bonanno, 2021).

In conclusion, the newly adapted FREE-Y, translated from the adult version, appears to be a psychometrically sound measure of EF in youth that can be recommended for wide use in both community and maltreated populations.

Acknowledgments

We gratefully acknowledge Keith F. Widaman for his statistical consultation. We also acknowledge Brenda Baney, Jennifer Craig, Gwen Miller, Paula Mulhall, Krista Rupert, and Amanda Thiess (Study 1); Kerri-Anne Bell, Katherine DiVasto, Olivia Pollak, and Katherine Tezanos (Study 2); and Alexis Broussard, Camila Caballero, and Sarah McCauley (Study 3) for assistance with data collection.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Research reported in this publication was supported by the National Institutes of Health (Sample 1: Eunice Kennedy Shriver National Institute of Child Health and Human Development under Award Number P50HD089922 and Sample 2: R15 MH113076-01A1). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ORCID iD

Ann-Christin Haag https://orcid.org/0000-0002-3865-8727

References

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review*, *30*(2), 217–237. https://doi.org/10.1016/j.cpr.2009.11.004

Aldao, A., Sheppes, G., & Gross, J. J. (2015). Emotion regulation flexibility. *Cognitive Therapy and Research*, *39*(3), 263–278. https://doi.org/10.1007/s10608-014-9662-4

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.

Assed, M. M., Khafif, T. C., Belizario, G. O., Fatorelli, R., Rocca, C. C., de A., & de Pádua Serafim, A. (2020). Facial emotion recognition in maltreated children: A systematic review. *Journal of Child and Family Studies*, *29*(5), 1493–1509. https://doi.org/10.1007/s10826-019-01636-w

Bardeen, J. R., Fergus, T. A., Hannan, S. M., & Orcutt, H. K. (2016). Addressing psychometric limitations of the difficulties in emotion regulation scale through item modification. *Journal of Personality Assessment*, *98*(3), 298–309. https://doi.org/10.1080/00223891.2015.1091774

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, *42*(5), 815–824. https://doi.org/10.1016/j.paid.2006.09.018

Benson, L., English, T., Conroy, D. E., Pincus, A. L., Gerstorf, D., & Ram, N. (2019). Age differences in emotion regulation strategy use, variability, and flexibility: An experience sampling approach. *Developmental Psychology*, *55*(9), 1951–1964. https://doi.org/10.1037/dev0000727

Berking, M., & Wupperman, P. (2012). Emotion regulation and mental health: Recent findings, current challenges, and future directions. *Current Opinion in Psychiatry*, *25*(2), 128–134. https://doi.org/10.1097/YCO.0b013e3283503669

Bernstein, I. H., Rush, A. J., Trivedi, M. H., Hughes, C. W., Macleod, L., Witte, B. P., Jain, S., Mayes, T. L., & Emslie, G. J. (2010). Psychometric properties of the Quick Inventory of Depressive Symptomatology in adolescents. *International Journal of Methods in Psychiatric Research*, *19*(4), 185–194. https://doi.org/10.1002/mpr.321

Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J., & Neer, S. M. K. (1997). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*(4), 545–553. https://doi.org/10.1097/00004583-199704000-00018

Bonanno, G. A. (2005). Resilience in the face of potential trauma. *Current Directions in Psychological Science*, *14*(3), 135–138.

Bonanno, G. A. (2021). The resilience paradox. *European Journal of Psychotraumatology*, *12*(1), Article 1942642. https://doi.org/10.1080/20008198.2021.1942642

Bonanno, G. A., & Burton, C. L. (2013). Regulatory flexibility: An individual differences perspective on coping and emotion regulation. *Perspectives on Psychological Science*, *8*(6), 591–612. https://doi.org/10.1177/1745691613504116

Bonanno, G. A., Colak, D. M., Keltner, D., Shiota, M. N., Papa, A., Noll, J. G., Putnam, F. W., & Trickett, P. K. (2007). Context matters: The benefits and costs of expressing positive emotion among survivors of childhood sexual abuse. *Emotion*, *7*(4), 824–837. https://doi.org/10.1037/1528-3542.7.4.824

Bonanno, G. A., Papa, A., Lalande, K., Westphal, M., & Coifman, K. (2004). The importance of being flexible: The ability to both enhance and suppress emotional expression predicts long-term adjustment. *Psychological Science*, *15*(7), 482–487. https://doi.org/10.1111/j.0956-7976.2004.00705.x

Burton, C. L., & Bonanno, G. A. (2016). Measuring ability to enhance and suppress emotional expression: The Flexible Regulation of Emotional Expression (FREE) Scale. *Psychological Assessment*, *28*(8), 929–941. https://doi.org/10.1037/pas0000231

Byrne, B. M. (2010). *Structural equation modeling with AMOS. Basic concepts, applications, and programming* (2nd ed.). Routledge.

Carlson, S. M., Koenig, M. A., & Harms, M. B. (2013). Theory of mind. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(4), 391–402. https://doi.org/10.1002/wcs.1232

Casey, B. J. (2015). Beyond simple models of self-control to circuit-based accounts of adolescent behavior. *Annual Review of Psychology*, *66*, 295–319. https://doi.org/10.1146/annurev-psych-010814-015156

Chang, C., Kaczkurkin, A. N., McLean, C. P., & Foa, E. B. (2018). Emotion regulation is associated with PTSD and depression among female adolescent survivors of childhood sexual abuse. *Psychological Trauma: Theory, Research, Practice, and Policy*, *10*(3), 319–326. https://doi.org/10.1037/tra0000306

Chaplin, T. M., & Aldao, A. (2013). Gender differences in emotion expression in children: A meta-analytic review. *Psychological Bulletin*, *139*(4), 735–765. https://doi.org/10.1037/a0030737

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Testing measurement invariance of second-order factor models. *Structural Equation Modeling*, *12*(3), 471–492.

Chen, S., Chen, T., & Bonanno, G. A. (2018). Expressive flexibility: Enhancement and suppression abilities differentially predict life satisfaction and psychopathology symptoms. *Personality and Individual Differences*, *126*, 78–84. https://doi.org/10.1016/j.paid.2018.01.010

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Cohen, A. O., Breiner, K., Steinberg, L., Bonnie, R. J., Scott, E. S., Taylor-Thompson, K., Rudolph, M. D., Chein, J., Richeson, J. A., Heller, A. S., Silverman, M. R., Dellarco, D. V., Fair, D. A., Galvan, A., & Casey, B. J. (2016). When is an adolescent an adult? Assessing cognitive control in emotional and non-emotional contexts. *Psychological Science*, *27*(4), 549–562. https://doi.org/10.1177/0956797617690343

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Coifman, K. G., & Bonanno, G. A. (2010). When distress does not become depression: Emotion context sensitivity and adjustment to bereavement. *Journal of Abnormal Psychology*, *119*(3), 479–490. https://doi.org/10.1037/a0020113

Coifman, K. G., & Summers, C. B. (2019). Understanding emotion inflexibility in risk for affective disease: Integrating current research and finding a path forward. *Frontiers in Psychology*, *10*, 392. https://doi.org/10.3389/fpsyg.2019.00392

Davis, E. L., Levine, L. J., Lench, H. C., & Quas, J. A. (2010). Metacognitive emotion regulation: Children's awareness that changing thoughts and goals can alleviate negative emotions. *Emotion*, *10*(4), 498–510. https://doi.org/10.1037/a0018428

Emery, L., & Hess, T. M. (2011). Cognitive consequences of expressive regulation in older adults. *Psychology and Aging*, *26*(2), 388–396. https://doi.org/10.1037/a0020041

Gioia, G., Isquith, P., Guy, S., & Kenworthy, L. (2015). *Behavior Rating Inventory of Executive Function (BRIEF-2)* (2nd ed.). Psychological Assessment Resources.

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. https://doi.org/10.1023/B:JOBA.0000007455.08539.94

Gupta, S., & Bonanno, G. A. (2011). Complicated grief and deficits in emotional expressive flexibility. *Journal of Abnormal Psychology*, *120*(3), 635–643. https://doi.org/10.1037/a0023541

Hollenstein, T., Granic, I., Stoolmiller, M., & Snyder, J. (2004). Rigidity in parent-child interactions and the development of externalizing and internalizing behavior in early childhood. *Journal of Abnormal Child Psychology*, *32*(6), 595–607.

Holodynski, M., & Friedlmeier, W. (2006). *Development of emotions and emotion regulation*. Springer.

Hourigan, S. E., Goodman, K. L., & Southam-Gerow, M. A. (2011). Discrepancies in parents' and children's reports of child emotion regulation. *Journal of Experimental Child Psychology*, *110*(2), 198–212. https://doi.org/10.1016/j.jecp.2011.03.002

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology*, *20*(1), 90–98. https://doi.org/10.1016/j.jcps.2009.09.003

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2021). *semTools: Useful tools for structural equation modeling* (R package version 0.5-4). https://CRAN.R-project.org/package=semTools

Kalokerinos, E. K., Greenaway, K. H., & Casey, J. P. (2017). Context shapes social judgments of positive emotion suppression and expression. *Emotion*, *17*(1), 169–186. https://doi.org/10.1037/emo0000222

Kim, J., & Cicchetti, D. (2010). Longitudinal pathways linking child maltreatment, emotion regulation, peer relations, and psychopathology. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *51*(6), 706–716. https://doi.org/10.1111/j.1469-7610.2009.02202.x

Kim-Spoon, J., Cicchetti, D., & Rogosch, F. A. (2013). A longitudinal study of emotion regulation, emotion lability-negativity, and internalizing symptomatology in maltreated and nonmaltreated children. *Child Development*, *84*(2), 512–527. https://doi.org/10.1111/j.1467-8624.2012.01857.x

Klimstra, T. A., Hale, W. W., Raaijmakers, Q. A. W., Branje, S. J. T., & Meeus, W. H. J. (2009). Maturation of personality in adolescence. *Journal of Personality and Social Psychology*, *96*(4), 898–912. https://doi.org/10.1037/a0014746

Kovacs, M. (2015). Children's Depression Inventory (CDI and CDI 2). In R. L. Cautin & S. O. Lilienfeld (Eds.), *The encyclopedia of clinical psychology* (pp. 1–5). Wiley-Blackwell.

Kromm, H., Färber, M., & Holodynski, M. (2015). Felt or false smiles? Volitional regulation of emotional expression in 4-, 6-, and 8-year-old children. *Child Development*, *86*(2), 579–597. https://doi.org/10.1111/cdev.12315

Langsrud, Ø. (2005). Rotation tests. *Statistics and Computing*, *15*, 53–60.

Langsrud, Ø., & Mevik, B.-H. (2019). *ffmanova: Fifty-Fifty MANOVA* (R package version 1.1.0). https://CRAN.R-project.org/package=ffmanova

Lovallo, W. R. (2013). Early life adversity reduces stress reactivity and enhances impulsive behavior: Implications for health behaviors. *International Journal of Psychophysiology*, *90*(1), 8–16. https://doi.org/10.1016/j.ijpsycho.2012.10.006

MacDermott, S. T., Gullone, E., Allen, J. S., King, N. J., & Tonge, B. (2010). The Emotion Regulation Index for Children and Adolescents (ERICA): A psychometric investigation. *Journal of Psychopathology and Behavioral Assessment*, *32*(3), 301–314. https://doi.org/10.1007/s10862-009-9154-0

Matthey, S., & Petrovski, P. (2002). The Children's Depression Inventory: Error in cutoff scores for screening purposes. *Psychological Assessment*, *14*(2), 146–149. https://doi.org/10.1037/1040-3590.14.2.146

Maydeu-Olivares, A., Shi, D., & Rosseel, Y. (2018). Assessing fit in structural equation models: A Monte-Carlo evaluation of RMSEA versus SRMR confidence intervals and tests of close fit. *Structural Equation Modeling*, *25*(3), 389–402. https://doi.org/10.1080/10705511.2017.1389611

Morris, A. S., Silk, J. S., Steinberg, L., Myers, S. S., & Robinson, L. R. (2007). The role of the family context in the development of emotion regulation. *Social Development*, *16*(2), 361–388. https://doi.org/10.1111/j.1467-9507.2007.00389.x

Nolen-Hoeksema, S. (2012). Emotion regulation and psychopathology: The role of gender. *Annual Review of Clinical Psychology*, *8*, 161–187. https://doi.org/10.1146/annurev-clinpsy-032511-143109

Noll, J. G. (2021). Child sexual abuse as a unique risk factor for the development of psychopathology: The compounded convergence of mechanisms. *Annual Review of Clinical Psychology*, *17*, 439–464. https://doi.org/10.1146/annurev-clinpsy-081219

Oshri, A., Kogan, S. M., Kwon, J. A., Wickrama, K. A. S., Vanderbroek, L., Palmer, A. A., & MacKillop, J. (2018). Impulsivity as a mechanism linking child abuse and

neglect with substance use in adolescence and adulthood. *Development and Psychopathology*, *30*(2), 417–435. https://doi.org/10.1017/S0954579417000943

Parsafar, P., Fontanilla, F. L., & Davis, E. L. (2019). Emotion regulation strategy flexibility in childhood: When do children switch between different strategies? *Journal of Experimental Child Psychology*, *183*, 1–18. https://doi.org/10.1016/j.jecp.2019.01.004

Pițur, S., & Miu, A. C. (2020). Childhood maltreatment and expressive flexibility: Specific effects of threat and deprivation? *Cognition and Emotion*, *34*(8), 1721–1728. https://doi.org/10.1080/02699931.2020.1795625

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Pynoos, R. S. S. A. M. (2015). *UCLA PTSD reaction index for children/adolescents- DSM-5*. Behavioral Health Innovations, LLC.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Rhemtulla, M., Brosseau-Liard, P. E., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*, 354–373. https://doi.org/10.1037/a0029315.supp

Richards, J. M., & Gross, J. J. (2006). Personality and emotional memory: How regulating emotion impairs memory for emotional events. *Journal of Research in Personality*, *40*(5), 631–651. https://doi.org/10.1016/j.jrp.2005.07.002

Rodin, R., Bonanno, G. A., Rahman, N., Kouri, N. A., Bryant, R. A., Marmar, C. R., & Brown, A. D. (2017). Expressive flexibility in combat veterans with posttraumatic stress disorder and depression. *Journal of Affective Disorders*, *207*, 236–241. https://doi.org/10.1016/j.jad.2016.09.027

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Rudnev, M., Lytkina, E., Davidov, E., Schmidt, P., & Zick, A. (2018). Testing measurement invariance for a second-order factor. A cross-national test of the alienation scale. *Methods, Data, Analyses*, *12*(1), 47–76. https://doi.org/10.12758/mda.2017.11

Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., Markowitz, J. C., Ninan, P. T., Kornstein, S., Manber, R., Thase, M. E., Kocsis, J. H., & Keller, M. B. (2003). The 16-item Quick Inventory of Depressive Symptomatology (QIDS), Clinician Rating (QIDS-C), and Self-Report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*, *54*, 573–583. https://doi.org/10.1016/S0006-3223(03)01866-8

Saarni, C. (1984). An observational study of children's attempts to monitor their expressive behavior. *Child Development*, *55*(4), 1504–1513.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507–514.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research*, *8*(2), 23–74.

Schrank, F., Mather, N., & McGrew, K. (2014). *Woodcock–Johnson IV tests of cognitive abilities examiner's manual, standard and extended batteries*. Riverside.

Schreier, H. M. C., Heim, C. M., Rose, E. J., Shalev, I., Shenk, C. E., & Noll, J. G. (2021). Assembling a cohort for in-depth, longitudinal assessments of the biological embedding of child maltreatment: Methods, complexities, and lessons learned. *Development and Psychopathology*, *33*(2), 394–408. https://doi.org/10.1017/S0954579420001510

Shields, A., & Cicchetti, D. (1998). Reactive aggression among maltreated children: The contributions of attention and emotion dysregulation. *Journal of Clinical Child Psychology*, *27*(4), 381–395. https://doi.org/10.1207/s15374424jccp2704_2

Silvers, J. A., McRae, K., Gabrieli, J. D. E., Gross, J. J., Remy, K. A., & Ochsner, K. N. (2012). Age-related differences in emotional reactivity, regulation, and rejection sensitivity in adolescence. *Emotion*, *12*(6), 1235–1247. https://doi.org/10.1037/a0028297

Skinner, E. A., & Zimmer-Gembeck, M. J. (2007). The development of coping. *Annual Review of Psychology*, *58*, 119–144. https://doi.org/10.1146/annurev.psych.58.110405.085705

Stifter, C., & Augustine, M. (2019). Emotion regulation. In V. LoBue, K. Pérez-Edgar, & K. A. Buss (Eds.), *Handbook of emotional development* (pp. 405–430). Springer Nature Switzerland. https://doi.org/10.1007/978-3-030-17332-6

Thibodeau, E. L., Cicchetti, D., & Rogosch, F. A. (2015). Child maltreatment, impulsivity, and antisocial behavior in African American children: Moderation effects from a cumulative dopaminergic gene index. *Development and Psychopathology*, *27*, 1621–1636. https://doi.org/10.1017/S095457941500098X

Van der Giessen, D., Hollenstein, T., Hale, W. W., Koot, H. M., Meeus, W., & Branje, S. (2015). Emotional variability in mother-adolescent conflict interactions and internalizing problems of mothers and adolescents: Dyadic and individual processes. *Journal of Abnormal Child Psychology*, *43*(2), 339–353. https://doi.org/10.1007/s10802-014-9910-9

Villalta, L., Smith, P., Hickin, N., & Stringaris, A. (2018). Emotion regulation difficulties in traumatized youth: A meta-analysis and conceptual review. *European Child and Adolescent Psychiatry*, *27*, 527–544. https://doi.org/10.1007/s00787-018-1105-4

Wang, Y., & Hawk, S. T. (2019). Expressive enhancement, suppression, and flexibility in childhood and adolescence: Longitudinal links with peer relations. *Emotion*, *20*(6), 1059–1073. https://doi.org/10.1037/emo0000615

Wang, Y., & Hawk, S. T. (2020). Development and validation of the Child and Adolescent Flexible Expressiveness (CAFE) Scale. *Psychological Assessment*, *32*(4), 358–373. https://doi.org/10.1037/pas0000795

Wang, Y., Hawk, S. T., & Zong, W. (2020). Bidirectional effects between expressive regulatory abilities and peer acceptance among Chinese adolescents. *Journal of Experimental Child Psychology*, *199*, 104891. https://doi.org/10.1016/j.jecp.2020.104891

Westphal, M., Seivert, N. H., & Bonanno, G. A. (2010). Expressive flexibility. *Emotion*, *10*(1), 92–100. https://doi.org/10.1037/a0018420

Zelazo, P. D. (2015). Executive function: Reflection, iterative reprocessing, complexity, and the developing brain. *Developmental Review*, *38*, 55–68. https://doi.org/10.1016/j.dr.2015.07.001

Zeman, J., Cassano, M., Perry-Parrish, C., & Stegall, S. (2006). Emotion regulation in children and adolescents. *Journal of Developmental and Behavioral Pediatrics*, *27*(2), 155–168. https://doi.org/10.1097/00004703-200604000-00014

Zeman, J., & Garber, J. (1996). Display rules for anger, sadness, and pain: It depends on who is watching. *Child Development*, *67*, 957–973.