

# Development and Validation of the Child and Adolescent Flexible Expressiveness (CAFE) Scale

Yingqian Wang and Skyler T. Hawk  
The Chinese University of Hong Kong

Expressive flexibility, or the ease with which one can switch between enhancing and suppressing emotional behavior in line with situational demands, appears to be an adaptive skill for promoting psychological well-being for adults. However, this construct has rarely been explored among children and adolescents. To facilitate such investigations, the present study developed and validated a scenario-based, self-reported measurement tool that tests youth's expressive flexibility, namely the Child and Adolescent Flexible Expressiveness (CAFE) Scale. Results from Study 1 ( $N = 549$ ) and Study 2 ( $N = 248$ ) supported a two-factor structure (expressive enhancement and expressive suppression), identifying eligible items through exploratory and confirmatory factor analyses. The measurement invariance of the scale was also evidenced via a series of multigroup tests across sex and age groups. Additionally, both subscales showed adequate internal consistencies, and the CAFE scores related to most theoretically similar measures (emotion expressivity and internal emotion regulation) and clinical outcome variables (e.g., depression, problem behavior), as expected. Study 3 ( $N = 199$ ) further compared CAFE scores to children's observed emotion expressivity in a standardized laboratory task. Results indicated that CAFE enhancement and suppression scores significantly predicted respective performances in the task. Finally, Study 4 ( $N = 48$ ) demonstrated adequate test-retest reliability of the CAFE Scale by retesting a subsample of participants from Study 1. Generally, the CAFE Scale appears to be a reliable and valid measurement in the area of emotion regulation, which allows convenient clinical assessment of children's and adolescents' expressive flexibility.

## **Public Significance Statement**

This article details the development and validation of the Child and Adolescent Flexible Expressiveness (CAFE) Scale, a new measurement designed to test youth's abilities to regulate emotional expressions as required by situational demands. Results of four studies supported the reliability and validity of the CAFE Scale, suggesting that it can be a useful tool for future research and clinical assessments related to emotion regulation.

**Keywords:** expressive flexibility, emotion regulation, scale development, children and adolescents

**Supplemental materials:** <http://dx.doi.org/10.1037/pas0000795.supp>

Although observable emotional expressions are commonly thought to reflect internal feelings, this is not always the case. According to a recent meta-analysis (Durán, Reisenzein, & Fernández-Dols, 2017), self-reported emotions and objectively assessed facial expressions are typically only moderately correlated. In many situations, individuals must avoid expressing their

internal feelings in order to conform to social expectations (e.g., restraining one's temper in public places), or must amplify emotional signals to send a message about what they want to happen (e.g., exaggerating sadness to gain sympathy). The ability to deploy these various expressive strategies in a flexible manner is an essential component of regulating emotions in line with contextual demands (Bonanno, Papa, Lalande, Westphal, & Coifman, 2004). Existing research has used both laboratory tasks (e.g., Westphal, Seivert, & Bonanno, 2010) and questionnaires (e.g., Chen, Chen, & Bonanno, 2018) to demonstrate the links between "expressive flexibility" and different aspects of adult psychological adjustment (e.g., less distress, higher life satisfaction, lower depression). To date, however, little is known about expressive flexibility among children and adolescents, and no measurement tool exists that specifically targets this group. Because childhood and adolescence are sensitive periods for the rapid development of a variety of social and emotional skills, it is of great interest to investigate the

This article was published Online First January 16, 2020.

Yingqian Wang and Skyler T. Hawk, Department of Educational Psychology, The Chinese University of Hong Kong.

The work described in this article was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project CUHK 14403514), awarded to Skyler T. Hawk.

Correspondence concerning this article should be addressed to Yingqian Wang, Department of Educational Psychology, The Chinese University of Hong Kong, Room 606, Chen Kou Bun Building, Shatin N.T., Hong Kong. E-mail: [wangyingqian@link.cuhk.edu.hk](mailto:wangyingqian@link.cuhk.edu.hk)

characteristics and functions of expressive flexibility at these stages. To facilitate such explorations, the present study aimed to develop a concise and convenient self-report measurement of expressive flexibility for youth, the Child and Adolescent Flexible Expressiveness (CAFE) Scale.

### Expressive Flexibility and Its Adaptive Value

The field of emotion regulation has blossomed over the last two decades. Many studies have examined various strategies to regulate internal feelings (see [Aldao, Nolen-Hoeksema, & Schweizer, 2010](#) for a review). The regulation of external expressions, however, has been relatively ignored. Definitions and theoretical models of emotion regulation have repeatedly emphasized the broad scope of regulatory behaviors, which could occur with any component of emotion (e.g., [Eisenberg & Spinrad, 2004](#); [Thompson, 1994](#)), including “internal feeling states, emotion-related physiological, attentional processes, motivational states, and/or the behavioral concomitants of emotion” ([Eisenberg & Spinrad, 2004](#), p. 338). Therefore, the regulation of overt, emotion-related behaviors is an important component of emotion regulation that is distinct from the regulation of subjective feelings and related cognitive or physiological processes. For example, efforts to reduce negative feelings versus negative expressions appear to activate different brain regions and yield discrepant neurological patterns (e.g., [Goldin, McRae, Ramel, & Gross, 2008](#)). [Zhu and Bonanno \(2017\)](#) have also shown that adults’ ability to up- and down-regulate subjective feelings was only moderately correlated with abilities in modulating expressive behaviors and that individuals are able to modify their emotional expressions independently of changes in their internal states. Moreover, both expressive and affective regulation abilities made unique contributions to predicting lower levels of depression. These findings suggest a need for instruments that assess individuals’ abilities to regulate their expressive behaviors, in addition to measures focused on the regulation of subjective states.

Within the realm of expressive regulation, enhancement and suppression are two related, but independent, methods of modulating emotional behaviors examined in many prior studies. Enhancement concerns exaggerating actual feelings or outwardly showing emotions that one is not actually experiencing; Suppression, on the other hand, refers to minimizing or hiding emotional expressions when necessary. Both strategies may result in desirable outcomes under certain situations, while creating problems under other conditions. For example, enhancing positive emotion can foster goal attainment in workplace interactions ([Wong, Tschan, Messerli, & Semmer, 2013](#)), but overly frequent amplification of positive emotions in daily life might actually reduce relationship quality due to feelings of inauthenticity ([Le & Impett, 2016](#)). Similarly, the ability to suppress overt expressions of pride might be socially desirable when outperforming others in one’s peer group ([Schall, Martiny, Goetz, & Hall, 2016](#), Study 3), but habitual suppression of both positive and negative emotions has been shown to be detrimental for both physical and psychological health (e.g., [Gross & John, 2003](#)). Instead of categorizing any one strategy as beneficial or costly, it is important to acknowledge that the rigid use of any particular strategy is likely to be maladaptive. To address this issue, [Bonanno et al. \(2004\)](#) proposed the notion of expressive flexibility, defined as the ability to switch between

enhancement and suppression of emotional expressions as required by situational demands. This construct contributes to the broader research on emotion regulation by focusing on the regulation of outward expressions and by considering the extent to which individuals can flexibly deploy diverse regulatory strategies.

According to the regulatory flexibility model ([Bonanno & Burton, 2013](#)), greater flexibility in expressive regulation requires context sensitivity (i.e., the ability to detect immediate demands of certain contexts and choose the most appropriate regulatory strategy), a diverse repertoire of strategies (the ability to utilize or enact various regulatory strategies that might fit with contextual demands), and feedback (the ability to monitor the effects of a chosen strategy and make proper adjustments). Prior studies of expressive flexibility have mainly focused on the repertoire component of this model, in terms of the ease with which individuals can engage in both expressive enhancement and suppression as required by their contexts. Thus, individuals are “flexible” if they can successfully both enhance and suppress their expressive behaviors and can easily switch from one strategy to the other.

Several studies have found evidence for the adaptiveness of expressive flexibility among adults. [Bonanno and colleagues \(2004\)](#) first explored this construct in a sample of New York City college students who had recently experienced the 9/11 terrorist attack. Results indicated that both enhancement and suppression abilities independently predicted less distress two years later, as did a flexibility score that considered the sum of these two separate abilities. [Westphal et al. \(2010\)](#) then replicated the positive link between expressive flexibility and friend-rated adjustment (indexed by mental health, physical health, social interactions, etc.) among individuals with high levels of life stress. [Gupta and Bonanno \(2011\)](#) compared the expressive flexibility scores of bereaved adults suffering from complicated grief disorder to both asymptomatic bereaved adults and to married adults. Results indicated that participants suffering from complicated grief after bereavement showed more deficits in expressive flexibility than the other two groups. Similarly, [Rodin and colleagues \(2017\)](#) showed that lower expressive flexibility predicted more severe posttraumatic stress and depressive symptoms among combat veterans. In a recent study, [Chen et al. \(2018\)](#) further demonstrated that the positive effects of expressive flexibility might also apply to non-Western cultures. In a sample of Chinese college students, expressive flexibility was positively correlated with resilience and life satisfaction, whereas negatively related with depression and anxiety. In summary, expressive flexibility appears to be an adaptive skill that is salutary for various aspects of individual mental health.

Because expressive enhancement and suppression constitute distinct abilities, several studies on the broader construct of expressive flexibility have also examined the unique contributions of these two component dimensions. In the study of [Rodin et al. \(2017\)](#), veterans’ enhancement ability was associated with lower levels of posttraumatic stress and depression, but suppression was not. The study by [Chen et al. \(2018\)](#) directly compared the predictive value of enhancement and suppression for indices of psychological adjustment. Results showed that enhancement ability uniquely predicted higher life satisfaction, while suppression ability significantly predicted lower depression and anxiety. These various findings suggest that enhancement and suppression are relatively independent processes that are sometimes associated

with different aspects of adjustment. Because of these unique associations and because overall scores of expressive flexibility are composited from scores of enhancement and suppression abilities, measures of these two constructs require separate examinations of their respective validity.

### Laboratory and Self-Report Measurements of Expressive Flexibility

Numerous instruments have been designed to measure the regulation of internal feelings, such as the Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004) and the Emotion Regulation Questionnaire (ERQ; Gross & John, 2003). In comparison, measurements designed to assess expressive regulation are scarcer, and might conflate knowledge of display rules with ability to enact those rules. Specific to expressive flexibility, a laboratory task (Bonanno et al., 2004) and a self-report scale (Burton & Bonanno, 2016) exist to assess this construct. Both measures were designed to examine the repertoire component of regulatory flexibility, namely the ability to both enhance and suppress expressions.

Bonanno and colleagues (2004) first developed the laboratory paradigm to measure expressive flexibility. In this task, participants view blocks of emotion-eliciting pictures and receive instructions to enhance expressions (even when they are not actually experiencing the emotion), suppress expressions, or behave normally. Blind coders then rate participants' overt emotional expressions. Thus, the "normal" condition provides a baseline measurement of expressivity and creates conditions for within-subject comparisons. Accordingly, enhancement ability is indexed by the expressive difference between enhancement and normal conditions, while suppression ability is indexed by the difference between normal and suppression conditions. Expressive flexibility is then computed by subtracting the polarity score of enhancement and suppression (i.e., the absolute value of the discrepancy between the two scores) from their sum (Westphal et al., 2010). This formula ensures that high expressive flexibility scores result from scoring high on both enhancement and suppression abilities. Generally, the laboratory paradigm provides an objective and rigorous measurement of expressive flexibility, and has been adopted by most previous studies examining this construct (Gupta & Bonanno, 2011; Rodin et al., 2017; Westphal et al., 2010). The benefit of this approach is that it assesses the relatively "pure" ability to enact these different strategies when instructed to do so, free of other social demands. As a laboratory task, however, it may fail to consider the influence of actual social contexts on expressive flexibility, and thus be deficient in terms of ecological validity (Burton & Bonanno, 2016). The complexity of conducting both the task and video coding also limits its feasibility in large-scale studies.

To address the limitations of laboratory assessment, Burton and Bonanno (2016) developed the self-reported Flexible Regulation of Emotional Expression (FREE) Scale, to measure expressive flexibility among adults. The scale consists of 16 scenarios, organized into four clusters based on the abilities they require: enhancing positive emotion, enhancing negative emotion, suppressing positive emotion, and suppressing negative emotion. The instructions before each cluster explicitly state the involved emotion (positive/negative) as well as the required ability (being more

expressive or suppressive of feelings). Participants report the degree to which they can modulate their expressions as required under each scenario from 1 = *not at all* to 6 = *very much*. Results indicated that both a four-factor model and a hierarchical model with enhancement and suppression abilities as second-order factors yielded good fits, but the internal consistencies of the two composite factors were higher (two composite factors:  $\alpha = .81$  and  $.70$ ; four separate factors:  $\alpha = .77, .65, .68,$  and  $.66$ ). Overall expressive flexibility and enhancement and suppression scores were similarly correlated with most personality and emotion measures, including higher use of cognitive reappraisal, agreeableness, emotional stability, openness, social functioning, and ego resilience, and lower affect regulation deficits and depression. Generally, expressive flexibility and its two subcomponents were all associated with better psychosocial adjustment.

Nevertheless, in line with the notion that enhancement and suppression are independent skills, these two dimensions still showed some differences in terms of their relationships with extraversion, conscientiousness, attitudes toward suppressing or expressing emotions, and, in particular, participants' habitual tendencies toward suppressing their emotions. Specifically, habitual suppression was negatively correlated with expressive enhancement ability, but showed a trend toward a positive correlation with expressive suppression ability. Additionally, both enhancement and suppression scores of the FREE Scale significantly predicted the corresponding behavioral performances in the laboratory task, indicating that individuals are capable of accurately assessing their own expressive flexibility. Chen and colleagues (2018) then adopted the FREE Scale in a study with Chinese college students, further demonstrating its adequate psychometric properties in a non-Western sample and reported positive links between self-reported expressive flexibility and psychological well-being.

Although the FREE Scale appears to be a reliable and valid scale for use with adult participants, it may not be applicable to children and adolescents. Most scenario items in this scale refer to experiences that frequently occur for adults but are not relevant for youth. Scenarios such as "A coworker gets a promotion and wants to talk about it" or "A friend is talking about a great date she had the other night" represent such problematic cases. Additionally, the original scale format asking individuals to assess "to what extent they could be able to modulate their expressions compared to how they were actually feeling" (Burton & Bonanno, 2016, p. 931) might be difficult for children to understand. Therefore, in order to extend expressive flexibility research to younger groups, a more concise self-report questionnaire targeted toward children and adolescents is needed.

### The Present Study

The present study aimed to develop the Child and Adolescent Flexible Expressiveness (CAFE) Scale as a convenient and reliable measurement of expressive flexibility for older children and adolescents. Following Bonanno's (Bonanno & Burton, 2013; Bonanno et al., 2004) definition of expressive flexibility, the CAFE Scale focuses on the ability to modulate emotional expressions synchronized with contextual demands, instead of the frequency of using certain regulatory strategies. In addition, to guarantee high ecological validity, all items of the CAFE Scale are based on real-life scenarios generated by children. We made some changes

in format, compared with the FREE Scale, in order to simplify it for child participants and to make items more straightforward. Specifically, each item contains both a scenario and a specific regulatory strategy. Therefore, children only needed to rate whether they could successfully enact the specific behavior described in each scenario.

In addition to a pilot study that collected scenarios for scale development, we conducted four studies to validate the new measure. In Studies 1 and 2, we tested the factor structure and reliability of the CAFE Scale, as well as the measurement equivalence across participant sex and age groups. We also explored potential sex and age group differences on the CAFE scores. Scale validity was assessed by charting the relationships between the CAFE Scale and a series of theoretically and clinically relevant variables. Expressive flexibility constitutes a component of emotion regulation that is distinct from the management of subjective states and related appraisals or cognitions. We therefore examined the unique predictive value of the CAFE Scale beyond existing emotion regulation measures, by supplementing the validity testing with a partial correlation analysis that controlled for internal emotion regulation difficulties. Study 3 further compared the CAFE scores with participants' observed performance in the laboratory task. Finally, Study 4 examined the test-retest reliability of the CAFE Scale.

### Scale Development

To ensure that the hypothetical scenarios used in the scale reflected youth's actual experiences, we first conducted a pilot study to collect possible scenarios. A total of 94 children in Grade 7 (51.1% female;  $M_{\text{age}} = 12.25$  years,  $SD = 0.53$ ) were recruited from a junior high school in Guangdong Province, China. In line with Bonanno et al.'s (2004) definition of expressive flexibility and the structure of the FREE Scale, two dimensions were designed for the CAFE Scale, namely Enhancement and Suppression. Therefore, participants in the pilot study recalled two kinds of situations they have experienced: (1) situations in which they needed to enhance/exaggerate their emotional expressions and (2) situations in which they needed to suppress/minimize their emotional expressions. Based on the face validity and the diversity of collected contexts (similar scenarios were merged), we selected 22 scenarios and reorganized them into scale items. To distinguish participants' perceived ability from their preference to use a specific regulatory strategy, we directly provided the desirable strategy (enhancement or suppression) in each item. Prior instructions also emphasized that the assessment on this questionnaire should be based on "whether you can do" but not "whether you are supposed to do" to ensure that the scale measured ability as opposed to knowledge.

There were 11 items in the Suppression subscale (e.g., "If someone makes me angry in a public place, I am able to control my impulse to cause a scene"; "When a classmate does something funny in class, I can refrain from laughing to help maintain order") and 11 items clustered into the Enhancement subscale (e.g., "I can laugh along when someone tells a joke that is not funny at all"; "If I have a great day at school but find my family member in a bad mood after going home, I can show empathy to him/her"). Thus, similar to those of the FREE Scale, Enhancement items included situations in which one would exaggerate a particular response even when actually feeling neutral or indifferent, as well as situ-

ations in which one would substitute or overlay a response that contradicted internal feelings. Although these latter situations could arguably include an element of suppression as well, the focus of the item remained on maximizing the display of a socially appropriate response. For this reason, and because children consistently generated these scenarios in response to the prompt about exaggerating emotional expressions, we classified such items as reflecting enhancement. The two subscales included both positive and negative emotions. Participants needed to rate their agreement to each statement on a five-point scale, ranging from 1 = *not at all true of me* to 5 = *extremely true of me*.

### Study 1: Item Selection and Initial Validation of the CAFE Scale

The aim of the first study was to determine the eligible items that could reflect expressive flexibility and provide initial validation about the psychometric properties of the CAFE Scale. Therefore, we first examined the new scale's factor structure through exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). We also examined measurement invariance via a set of model comparisons between boys and girls, as well as between primary school and junior high school students. Next, the internal consistencies of the two subscales were respectively calculated, followed by an initial exploration of sex and age differences on CAFE scores. Finally, the validity of CAFE Scale was tested by charting its relationships with preexisting measures, including the relevant constructs of emotional expressivity and internal emotion regulation difficulties (for testing convergent validity) as well as clinically emotional, behavioral, and social outcome variables (for testing criterion validity). In order to examine the unique contributions of expressive flexibility in predicting individual adjustment, over and above scores of internal emotion regulation abilities, we additionally calculated the partial correlations between CAFE scores and outcome variables when controlling for self-reported difficulties in internal emotion regulation.

Regarding the scale structure, we expected to find the enhancement and suppression factors reported in prior research. Regarding validity testing, our predictions were as follows: First, because the two subscales of the CAFE respectively measured the ability to modulate emotional expressions upward and downward, it was hypothesized that the emotional expressivity score would correlate positively with the CAFE enhancement score and correlate negatively with the CAFE suppression score. Second, expressive flexibility and classically defined emotion regulation reflect the modulation of external expression and internal experience, respectively, but both belong to the larger domain of self-regulation. Therefore, we expected that higher CAFE scores would relate to less difficulty in internal emotion regulation. Third, as previous studies have established expressive flexibility as an adaptive ability (see Bonanno & Burton, 2013 for a review), we anticipated that the expressive flexibility and two subscale scores would all significantly relate with better emotional, behavioral, and social functioning, even after controlling for internal emotion regulation.

### Participants

Participants ( $N = 570$ ) were recruited from one primary school and one junior high school in the rural area of Shandong Province,

China. All participants were voluntary with parental or guardian permission. Data of 21 children were excluded because of obviously random responding. Thus, there remained 549 valid participants (49.4% female), aged between 8 and 16 ( $M_{\text{age}} = 12.42$  years,  $SD = 1.70$ ). They were respectively from Grade 4 ( $N = 90$ ;  $M_{\text{age}} = 10.13$  years,  $SD = 0.80$ ), Grade 5 ( $N = 95$ ;  $M_{\text{age}} = 10.92$  years,  $SD = 0.36$ ), and Grade 6 ( $N = 96$ ;  $M_{\text{age}} = 11.91$  years,  $SD = 0.35$ ) of primary school, as well as Grade 7 ( $N = 93$ ;  $M_{\text{age}} = 12.87$  years,  $SD = 0.55$ ), Grade 8 ( $N = 85$ ;  $M_{\text{age}} = 13.82$  years,  $SD = 0.37$ ), and Grade 9 ( $N = 90$ ;  $M_{\text{age}} = 14.91$  years,  $SD = 0.34$ ) of junior high school.

## Measures

A bilingual researcher, whose first language was Chinese, translated all measures without existing Chinese versions into Chinese. To ensure accuracy, another bilingual research assistant, whose first language was English, then conducted a back translation.

### Convergent validity measures.

**Emotional Expressivity Scale (EES).** The EES (Kring, Smith, & Neale, 1994) is a 17-item, self-reported questionnaire that assesses general emotional expressivity. It consists of six positively scored items and 11 reverse-scored items. Concerning children's limited understanding ability, we only used six positive-scored items to make the report easier (e.g., "I think of myself as emotionally expressive") in the present study. Additionally, to keep consistent with other emotion measures, we revised the original six-point Likert scale into a five-point scale (1 = *not at all true of me* to 5 = *extremely true of me*). The average score on all items was calculated, with high scores representing greater emotional expressivity. The Cronbach's alpha was .59 in the present study.

**Difficulties in Emotion Regulation Scale (DERS).** The DERS (Gratz & Roemer, 2004) is an integrative emotion regulation measurement. The present study selected three dimensions that closely relate to the regulatory ability of internal emotional experience, including Limited Access to Effective Emotion Regulation Strategies (e.g., "When I am upset, I believe that wallowing in it is all I can do"), Difficulties Engaging in Goal-Directed Behavior (e.g., "When I am upset, I have difficulty focusing on other things"), and Impulse Control Difficulties (e.g., "When I am upset, I lose control over my behaviors"). A total of 19 items of the three dimensions were rated from 1 = *almost never* to 5 = *almost always*, with higher scores indicating greater difficulties on emotion regulation. The internal consistencies of each dimension and the whole scale were  $\alpha = .71, .75, .77$ , and  $.88$ , respectively.

### Criterion validity measures: Psychopathology.

**Child Depression Inventory (CDI).** Children's depressive symptomatology was measured with the CDI (Kovacs & Beck, 1977), which was specifically designed for youth between 7 to 18 years. The scale comprises 27 items with three-point answers (e.g., 0 = *I am sad once in a while* to 2 = *I am sad all the time*). Participants responded to each item according to their state in the last 2 weeks, with higher average score reflecting more depressive symptoms. The Cronbach's alpha was .84 in the present study.

**Affect Balance Scale (ABS).** The ABS (Bradburn, 1969) is a 14-item questionnaire assessing people's daily affective state. It consists of eight items for Positive Affect (e.g., "things are going your way") and six items for Negative Affect (e.g., "very lonely or remote from other people"). Participants were asked to rate each

item according to the frequency of each experience in the last 1 month on a four-point scale (1 = *never* to 4 = *always*). The Cronbach's alpha of two subscales were .80 and .73, respectively.

**Problem Behavior Frequency Scale (PBFS).** Participants' externalizing problems were measured with the PBFS (Farrell, Kung, White, & Valois, 2000), which comprises four dimensions: Delinquency (six items; e.g., "cheated on a test"), Physical Aggression (six items; e.g., "hit or slapped another kid"), Nonphysical Aggression (seven items; e.g., "spread a rumor"), and Drug Use (five items; e.g., "used marijuana"). Because drug use is very uncommon among Chinese children and adolescents, we did not use this dimension. Participants were asked to honestly rate the frequency of having each behavior in the last half year on a five-point scale ranging from 0 *times* to 6 *times or more*. The internal consistencies of three subscales and the whole scale in the present study were  $\alpha = .79, .80, .86$ , and  $.92$ , respectively.

### Criterion validity measures: Social functioning.

**Social Peer Rejection Measure (SPRM).** The SPRM (Lev-Wiesel, Sarid, & Sternberg, 2013) was used to assess participants' self-perceived peer rejection. The scale includes 21 items and four dimensions: Insult (e.g., "I was called names"), Ignore (e.g., "I was rejected by my class/group"), Accusation (e.g., "Friends spread negative rumors about me"), Physical Attacks and Bossiness (e.g., "I had objects thrown at me by friends"). Response options range from 1 = *never happened* to 5 = *happened all the time*. The internal consistencies of four subscales and the whole scale in the present study were  $\alpha = .88, .87, .79, .79$ , and  $.94$ , respectively.

**Nominated peer status.** Besides participants' self-reported social functioning, we also measured their objective peer status with peer nomination techniques. Participants were asked to circle the names of three classmates they liked most on a class list and three whom they liked least. The numbers of nominations that children received on the two items, defined as Peer Acceptance and Peer Rejection, respectively, were first standardized within class to  $M = 0$  and  $SD = 1$ . Peer Status was then computed by subtracting the Peer Rejection score from the Peer Acceptance score for each child (Cillessen & Bukowski, 2000).

## Procedure

The Human Ethics Review Board of The Chinese University of Hong Kong approved all procedures. Participants and their parents received information about the study, including research aims, procedures, risks/benefits, the voluntary nature of participation, and right to withdraw. This is also the case for all other studies reported in this article. Children who agreed to participate were group tested in quiet classrooms and completed a series of questionnaires in a 40-min session. A trained instructor explained the requirements before start and supervised the whole process. Upon completion, participants received small gifts for their participation.

## Results

### Factor Analyses

Following the guidelines for validating a newly developed measurement (Worthington & Whittaker, 2006), we began with EFA to explore the underlying factor structure of the CAFE Scale, followed by CFA to evaluate the solution obtained from EFA. To

achieve this cross-validation, we randomly divided the whole sample into roughly two equal parts (Anderson & Gerbing, 1988): Sample A ( $N = 265$ ; 50.4% female;  $M_{\text{age}} = 12.46$  years,  $SD = 1.70$ ) was used for EFA, whereas Sample B ( $N = 284$ ; 48.6% female;  $M_{\text{age}} = 12.39$  years,  $SD = 1.71$ ) for CFA.

**Exploratory factor analysis.** We first inspected the correlation matrix of all items. Four items (Items 1, 12, 18, 22) were excluded from further analyses because of low correlations with other items (more than half of 21 correlation coefficients were not significant). Responses to the remaining 18 items were subjected to EFA with Varimax rotation and a fixed factor number of two (reflecting Enhancement and Suppression). Regarding item selection, factor loadings of .40 and higher were considered as meaningful. In addition, items that had double loadings or those that loaded on the unintended factor were also excluded. Following these criteria, five items were deleted after the first EFA: Items 5, 7, 10, and 19 had factor loadings below .40 on both dimensions; Item 3, which was intended as a Suppression item, loaded on the Enhancement factor. We reran the EFA on the remaining 13 items to ensure that all had satisfactory factor loadings (see Table 1 for final items and their factor loadings). Upon extraction and rotation, the two factors accounted for 41.11% of the total variance. The eigenvalue of the Enhancement factor was 4.02, which explained 21.71% of the variance and consisted of seven items. The eigenvalue of the Suppression factor was 1.32, which explained 19.40% of the variance and consisted of six items.

**Confirmatory factor analysis.** We conducted CFA with Mplus 7.0 (Muthén & Muthén, 1998–2017) to further assess the fitness of the measurement model using Sample B. The Shapiro-Wilk test indicated that scores for many items were not normally distributed (Shapiro-Wilk values ranging between .79 and .89,

$ps < .001$ ). Thus, we used the robust maximum likelihood (MLR) estimation in CFA and the following invariance tests, which could yield parameter estimates and a Satorra-Bentler chi-square statistic ( $\chi^2_{\text{SB}}$ ) that are robust to nonnormality (Satorra & Bentler, 1994). The model fit was considered acceptable when the comparative fit index (CFI) and Tucker-Lewis index (TLI) values were at or above .90, root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) values were at or below .08 (Kline, 2011). Results indicated that all indexes of the measurement model fell into acceptable range:  $\chi^2_{\text{SB}}(64) = 93.30$ ,  $p = .010$ , CFI = .94, TLI = .93, RMSEA = .04, SRMR = .05, which further validated the two-factor structure of the CAFE Scale. The factor loadings of seven Enhancement items ranged from .42 to .64, whereas that of six Suppression items were between .47 and .70, all of which were beyond the critical value of .40 (see Figure 1).

### Measurement Invariance

We then examined the measurement invariance of the CAFE Scale through a series of multigroup tests. Comparisons were made between boys and girls, as well as between primary and junior high school participants. Following the independent structure replications with the two split samples, we conducted these tests with the full sample in order to ensure adequate sample sizes in each subgroup. For each comparison, we constructed four sequentially nested models. In the baseline model (M0), all parameters were freely estimated between groups. The next three models sequentially added equal constraints across groups on factor loadings (M1), latent factor variances (M2), and latent factor covariances (M3). Therefore, four levels of measurement invariance were

Table 1  
Factor Loadings, Means, and Standard Deviations of the CAFE Items (Study 1)

CAFE item	Loading		<i>M</i>	<i>SD</i>
	Factor 1	Factor 2		
<b>Enhancement subscale</b>				
Item 11: If someone spends a lot of effort to cook a meal but it doesn't taste very good, I am able to pretend to enjoy it.	.75	.14	3.80	1.19
Item 20: If a friend beats me in a competition, I am able to appear happy for him/her no matter how disappointed I feel.	.69	.17	3.57	1.20
Item 21: When hanging out with friends at a place that I don't like, I can still make it seem like I'm having a good time.	.69	.07	3.35	1.18
Item 9: When I get a gift that I don't like very much, I can still act happy and grateful to avoid hurting others' feelings.	.60	.22	4.04	1.12
Item 17: If a friend of mine has a certain hobby (such as singing and painting), I am good at praising his/her works even I think it is just average.	.53	.27	3.47	1.19
Item 16: I can laugh along when someone tells a joke that is not funny at all.	.46	.07	2.83	1.29
Item 4: If I have a great day at school but find my family member in a bad mood after going home, I can restrain my joy and show empathy to him/her.	.44	.29	3.49	1.26
<b>Suppression subscale</b>				
Item 2: If I win the first prize in a competition, I can act humble and modest.	.05	.74	3.45	1.12
Item 15: When I get a good grade on an exam, I can remain outwardly calm so that others who did poorly won't feel bad.	.11	.69	3.63	1.14
Item 14: If someone makes me angry in a public place, I am able to control my impulse to cause a scene.	.23	.62	3.83	1.15
Item 13: I can control my anger in an argument to prevent things from becoming worse.	.23	.61	3.40	1.19
Item 6: When a classmate does something funny in class, I can refrain from laughing to help maintain order.	.16	.53	3.28	1.30
Item 8: When a classmate that I don't like comes to talk to me, I am still able to hide my dislike and carry on a normal conversation.	.36	.42	3.58	1.19

Note. CAFE = Child and Adolescent Flexible Expressiveness Scale.

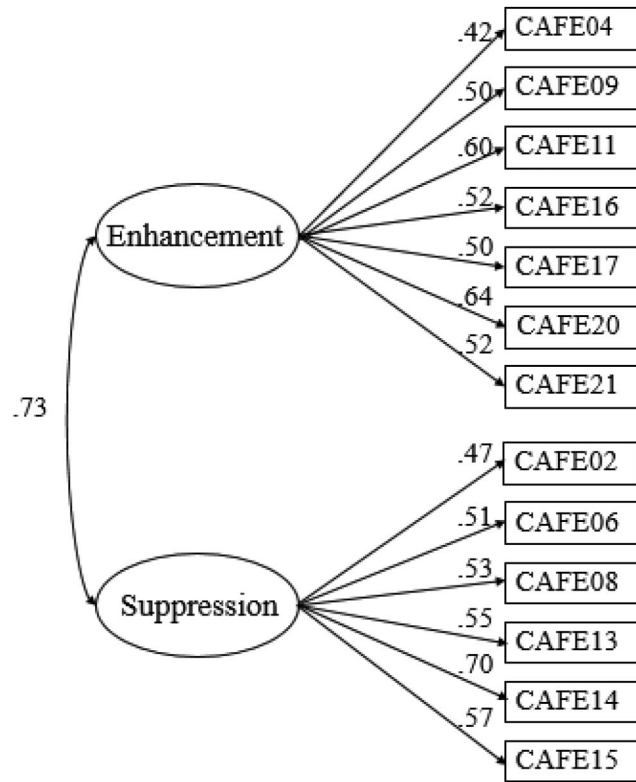


Figure 1. Standardized factor loadings of the Child and Adolescent Flexible Expressiveness (CAFE) Scale according to confirmatory factor analysis results (Study 1).

examined: Configural invariance was indexed by the model fits of unconstrained M0; Metric invariance was tested by comparing M1 with M0; Similarly, latent factor variance invariance and latent factor covariance invariance were respectively tested by comparing M2 with M1, and M3 with M2 (Cheung & Rensvold, 2002). For each model comparison, we computed the adjusted chi-square difference, with a significant change indicating noninvariance (Satorra & Bentler, 2001). However, this test might be overly sensitive for the current sample size (>200). We therefore regarded a decrease in CFI  $\leq$  .010, supplemented by an increase in RMSEA  $\leq$  .015 as more informative indicators of noninvariance (Chen, 2007; Hawk et al., 2013).

Results of the invariance tests are presented in Table 2. For both sex and age group comparisons, the baseline models showed adequate fits to the data (CFIs = .93 and .91, RMSEAs = .04 and .05, SRMRs = .05 and .06, respectively), which demonstrated the configural invariance of the CAFE Scale. Regarding the following model comparisons, most chi-square differences were nonsignificant ( $ps \geq .237$ ), with the only exception being the M2–M1 comparison for different age groups ( $p = .005$ ). Nevertheless, none of the subsequent constraints yielded  $\Delta CFI > .010$  or  $\Delta RMSEA > .015$ , so the scale could be considered as invariant across sex and age groups at the metric, factor variance, and factor invariance levels.

**Reliability**

Reliability and validity analyses were also based on the whole sample. Cronbach’s alpha was calculated to determine the internal consistency of the CAFE items. Results indicated that both Enhancement and Suppression subscales had adequate internal consistencies, with  $\alpha = .73$  and  $.72$ , respectively. Item-total correlations (correlations between specific item scores and the average score of their respective dimension) were all significant, ranging between .57 to .69.

**Exploration of Sex and Age Group Differences**

We additionally examined the sex and age group differences on Expressive Flexibility, Enhancement, and Suppression, through a 2 (sex: girls vs. boys)  $\times$  2 (age group: primary school students vs. junior high school students) between-subjects multivariate analysis of variance. CAFE Enhancement and Suppression scores were calculated by averaging corresponding items within each subscale; the overall Expressive Flexibility score was calculated by subtracting the absolute value of the difference between the Enhancement and Suppression scores from their sum, that is, Expressive Flexibility = (Enhancement + Suppression) – |Enhancement – Suppression| (Westphal et al., 2010). None of the CAFE scores yielded a significant interaction between sex and age group ( $ps \geq .251$ ). The main effect of sex was significant on the Suppression score,  $F(1, 540) = 7.56, p = .006, \eta^2 = .01$ , with girls ( $M = 3.63, SD = 0.78$ ) scoring significantly higher than boys ( $M = 3.46, SD = 0.79$ ). However, there were no significant sex differences on either Expressive Flexibility or Enhancement scores ( $ps = .210$  and  $.641$ ). On the other hand, results revealed significant age group

Table 2  
Multigroup Tests of Measurement Invariance (Study 1)

Comparison	Model	$\chi^2_{SB}$	df	CFI	RMSEA	SRMR	Model	Adjusted $\Delta\chi^2_{SB}$	$\Delta CFI$	$\Delta RMSEA$
Male vs. female	M0. Configural invariance	205.023	139	.934	.042	.051				
	M1. Full metric invariance	218.956	150	.932	.041	.060	M1–M0	13.809	–.002	–.001
	M2. Factor variance invariance	219.923	152	.933	.041	.060	M2–M1	.334	.001	.000
	M3. Factor covariance invariance	220.009	153	.934	.040	.060	M3–M2	.086	.001	–.001
Primary vs. junior high	M0. Configural invariance	224.777	139	.910	.047	.056				
	M1. Full metric invariance	238.430	150	.907	.046	.063	M1–M0	13.932	–.003	–.001
	M2. Factor variance invariance	246.182	152	.901	.048	.079	M2–M1	10.431**	–.006	.002
	M3. Factor covariance invariance	246.411	153	.902	.047	.080	M3–M2	.031	.001	–.001

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; SB = Satorra-Bentler.

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

differences on all the three scores; Expressive Flexibility:  $F(1, 540) = 24.51, p < .001, \eta^2 = .04$ ; Enhancement:  $F(1, 540) = 11.94, p = .001, \eta^2 = .02$ ; Suppression:  $F(1, 540) = 44.69, p < .001, \eta^2 = .08$ . Specifically, primary school participants reported significantly higher Expressive Flexibility (primary:  $M = 6.81, SD = 1.60$ ; junior high:  $M = 6.20, SD = 1.25$ ), Enhancement (primary:  $M = 3.67, SD = 0.81$ ; junior high:  $M = 3.45, SD = 0.68$ ), and Suppression (primary:  $M = 3.76, SD = 0.81$ ; junior high:  $M = 3.33, SD = 0.70$ ) scores, compared with junior high school participants.

## Validity

Correlations between the CAFE Scale and other validity measures are depicted in Table 3.

**Convergent validity.** Convergent validity was analyzed by examining relations between the CAFE Scale and two relevant emotion measures, namely the External Emotional Expressivity Scale (EES) and the Difficulties in (Internal) Emotion Regulation Scale (DERS). According to the results, participants' Emotional Expressivity showed a trend toward a positive correlation with CAFE Enhancement score ( $p = .059$ ) but was not correlated with CAFE Suppression score or Expressive Flexibility score. Therefore, consistent with our expectation, children reporting higher Enhancement score tended to be more habitually expressive. However, there was not a negative correlation between Emotional

Expressivity and CAFE Suppression, which contradicted our hypothesis. Additionally, the overall Expressive Flexibility score was negatively correlated with the average score of DERS and its Limited Access to Effective Emotion Regulation Strategies, Difficulties Engaging in Goal-Directed Behavior, and Impulse Control Difficulties dimension scores. Thus, the more children viewed themselves as lacking expressive flexibility, the more difficulties they had with internal emotion regulation. CAFE Enhancement and Suppression scores similarly showed negative correlations with the overall DERS and subscale scores, though the link between CAFE Enhancement and DERS Strategies was a trend ( $p = .054$ ).

**Criterion-related validity.** Criterion-related validity was assessed by examining the extent to which the CAFE Scale was associated with clinically relevant psychological, behavioral, and social outcomes. We first examined associations with internalizing emotional states, including Depression (CDI) and Affect Balance (ABS). Expressive Flexibility, Enhancement, and Suppression scores were all negatively correlated with Depression and positively correlated with Positive Affect; Suppression was also negatively correlated with Negative Affect, whereas overall Expressive Flexibility showed a trend toward negative link with Negative Affect ( $p = .076$ ). Regarding the Externalizing Behaviors measured by the PBFS, overall Expressive Flexibility and its two subscale scores all showed significant negative correlations with

Table 3

Measures of Validity: Means, Standard Deviations, and Correlations With CAFE Enhancement, Suppression, and Expressive Flexibility Scores (Study 1)

Measure	<i>M</i>	<i>SD</i>	Bivariate correlations			Partial correlations after controlling the DERS score		
			Enhancement	Suppression	Expressive flexibility	Enhancement	Suppression	Expressive flexibility
Convergent validity measures								
Emotional expressivity	2.47	.67	.08 <sup>†</sup>	-.03	-.01			
Difficulties in emotion regulation								
Impulse	2.15	.78	-.19***	-.22***	-.22***			
Goals	2.61	.88	-.09*	-.21***	-.16***			
Strategies	2.35	.69	-.08 <sup>†</sup>	-.15***	-.14**			
Overall	2.36	.66	-.14**	-.22***	-.20***			
Criterion validity measures: Psychopathology								
Depression	1.49	.26	-.17***	-.25***	-.25***	-.12**	-.15**	-.17***
Affect balance								
Positive affect	2.93	.64	.19***	.21***	.22***	.16***	.15**	.18***
Negative affect	2.00	.57	-.03	-.11*	-.08 <sup>†</sup>	.06	-.01	.03
Problem behavior								
Delinquency	1.15	.41	-.16***	-.20***	-.21***	-.17***	-.20***	-.22***
Physical aggression	1.25	.49	-.15***	-.27***	-.25***	-.15**	-.24***	-.24***
Nonphysical aggression	1.20	.46	-.18***	-.22***	-.22***	-.19***	-.20***	-.22***
Overall	1.20	.43	-.17***	-.25***	-.24***	-.18***	-.23***	-.24***
Criterion validity measures: Social functioning								
Social peer rejection								
Insult	1.66	.79	-.05	-.12**	-.14**	.00	-.03	-.06
Ignore	1.66	.65	-.04	-.08 <sup>†</sup>	-.08 <sup>†</sup>	.01	.03	.01
Accusation	1.70	.76	-.11*	-.15***	-.16***	-.06	-.09 <sup>†</sup>	-.11*
Attack	1.54	.71	-.10*	-.10*	-.15**	-.06	.01	-.07
Overall	1.64	.64	-.08 <sup>†</sup>	-.13**	-.14**	-.02	-.02	-.06
Peer status	.00	1.59	.07	.11*	.10*	.05	.07	.06

Note. CAFE = Child and Adolescent Flexible Expressiveness Scale; DERS = Difficulties in Emotion Regulation Scale.

<sup>†</sup>  $p < .08$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .



each dimension of the PBFS. In other words, children who had greater flexibility on emotional expressions reported lower problem behaviors such as Delinquency and Aggression. Of the social measures, Expressive Flexibility and Suppression scores were negatively correlated with nearly all aspects of self-reported Peer Rejection (except the Ignore dimension, which trended toward significant links with Expressive Flexibility and Suppression at  $p = .055$  and  $p = .065$ ) and were positively correlated with nominated Peer Status. The Enhancement score, in contrast, was only negatively correlated with two dimensions of Peer Rejection (Accusation and Attack) and showed a trend toward negative correlation with overall Peer Rejection ( $p = .073$ ).

We also conducted partial correlation analyses, controlling for internal emotion regulation problems, as indexed by the DERS overall score. As is shown in Table 3, most correlations between CAFE scores and outcome variables remained significant. Expressive Flexibility and its two subscale scores retained significant associations with lower Depression, higher Positive Affect, and fewer Problem Behaviors. Although the previous modest correlations between CAFE scores and social functioning measures became nonsignificant, the changes in correlation coefficients were minimal. These results suggest that our measure of expressive flexibility still significantly predicted youth's psychological well-being, over and above a widely utilized index of internal emotion regulation difficulties.

## Discussion

The findings from this study lent initial support to the CAFE Scale as a reliable and valid measure of expressive flexibility in children and adolescents. The results of factor analyses showed an acceptable fit for the two-factor structure. Thus, the CAFE Scale was appropriate for its intended goal: to measure expressive flexibility via its two component dimensions, expressive enhancement and suppression abilities. The item loadings on their respective factors were generally acceptable. Both subscales had good internal consistencies.

A series of multigroup tests evidenced the measurement invariance of the CAFE Scale between boys and girls and between primary and junior high school students. In other words, the scale measured the same construct across different sex and age groups, suggesting that the scores of different groups are directly comparable. Therefore, following the invariance test, we additionally explored sex and age differences in CAFE scores. Results indicated that girls scored significantly higher than boys on suppression ability but not on expressive flexibility or enhancement ability. In line with traditional gender norms, girls are more likely to be socialized to hide high-intensity emotions (Chaplin, Cole, & Zahn-Waxler, 2005), which might result in greater suppression abilities. This might be especially the case in collectivist cultural contexts, where emotion suppression is valued for promoting in-group harmony (Matsumoto, Yoo, Nakagawa, & Multinational Study of Cultural Display Rules, 2008). Regarding age differences, junior high school students scored lower on all aspects of the CAFE Scale compared with primary school students. Although this seems to contradict the conventional wisdom that children's self-regulatory abilities increase with age, such development is not necessarily linear. Several previous studies have revealed that the development of internal emotion regulation (Zimmermann &

Iwanski, 2014) and emotional clarity (Haas et al., 2019) also show a dip in early adolescence. Indeed, early adolescence is characterized by both neurological changes and novel social and emotional experiences (Crone & Dahl, 2012) that may result in temporary stagnancy in the development of emotion-related abilities, including expressive flexibility. Additionally, younger children's confidence and optimism might also lead to their higher scores on this self-report measurement. These diverse explanations point to the need for more detailed research on age and sex differences, especially studies utilizing longitudinal or (quasi-)experimental designs.

Expressive flexibility and its two dimension scores showed modest negative relationships with difficulties in internal emotion regulation. This finding is consistent with the previous validation results for the FREE Scale and suggests that CAFE Scale could reflect children's regulatory abilities. Conversely, the two subscales showed slightly different correlation patterns with habitual emotional expressivity (only CAFE enhancement score showed a positive correlation), which provided initial support for the differences between enhancement and suppression. Although we anticipated that there would be a negative association between the CAFE suppression score and habitual expressivity, the absence of a significant relationship is also understandable, because both the EES and CAFE enhancement dimension measure the degree of showing emotions, which is not necessarily diametrically opposed to the degree that children are able to hide their emotions. However, the correlation coefficient between CAFE enhancement score and expressivity score only trended toward significance, which may be partially due to the poor reliability of the EES. Therefore, further evidence is needed to show that the two CAFE subscales reflect distinct abilities.

In terms of the criterion-related validity, the expressive flexibility and suppression scores of the CAFE Scale significantly predicted almost all indices of psychological, behavioral, and social adjustment, including lower depression, higher positive affect, lower negative affect, less problem behaviors, lower rejection from peers, and greater peer status. Nevertheless, the CAFE enhancement score was more variable in its relations with the adjustment variables, holding significant links with depression, positive affect, problem behavior, and two of the four peer rejection dimensions, but not with peer status or negative affect. This is potentially related to the collectivistic cultural context of the present study, in which it is highly encouraged to inhibit emotion expressions in order to promote in-group harmony (Matsumoto et al., 2008). Accordingly, expressive suppression may be more favored in social interactions. This might also be why the social measures, in particular, showed more consistent relations with suppression than with enhancement. In general, however, expressive flexibility and its two subcomponents all contributed to more positive emotional states, less problem behaviors, and better social relationships. The additional partial correlation analysis demonstrated that, even after controlling for difficulties in internal emotion regulation, most of the aforementioned correlations remained significant. This suggests that the CAFE Scale held unique predictive value for children's psychological and behavioral functioning, beyond the variance explained by the DERS. Thus, the DERS and the CAFE Scale appear to assess distinct sets of abilities that fall under the broader construct of emotion regulation, and both internal and external

emotion regulation might hold unique relationships with youth's psychological adjustment.

### Study 2: Further Validation of the CAFE Scale

Although Study 1 provided initial evidence that the CAFE Scale is a reliable measurement tool, two main limitations still existed. First, the sample was not representative enough. Participants all came from rural schools, which might restrict the generalizability of the new scale. Second, Study 1 yielded little evidence of differences between the enhancement and suppression subscales, indicating a further need for evidence of discriminant validity. Therefore, we conducted another study to further investigate the properties of the CAFE Scale using an urban sample. In Study 2, we first examined whether the factor structure and scale reliabilities could be replicated in this new sample. We retained emotion expressiveness as an index of convergent validity, since it is a very close notion with expressive flexibility and theoretically should show different correlation patterns with CAFE enhancement and suppression abilities. However, the poor reliability of EES in Study 1 suggested that it might not be a good candidate for further inclusion. This time we adopted two additional emotional expressivity scales and differentiated between positive and negative expressivity. We anticipated that positive expressivity would be more strongly correlated with enhancement ability than with suppression ability. Nevertheless, overly frequent expression of negative emotion may be generally perceived as maladaptive, particularly in the Chinese context (Trommsdorff & Rothbaum, 2009), so we anticipated that it might negatively correlate with expressive flexibility and both subscales scores. We also tested Big Five personality traits as another index to differentiate between expressive enhancement and suppression. In line with the results of the FREE Scale validation study, we hypothesized that the main difference would manifest in terms of extraversion, which might only positively correlate with enhancement ability. Finally, we included ego resilience as a measure of psychological flexibility and expected positive correlations with overall expressive flexibility and the two subscales.

### Participants and Procedure

We recruited 248 participants from a junior high school in urban area of Shandong Province, China. Participants were all in Grade 7, aged between 12 to 14 years ( $M_{\text{age}} = 12.74$  years,  $SD = 0.38$ ). The proportion of female was slightly higher, at 51.2%. Participants completed the CAFE Scale and several other measures under the supervision of a trained instructor, upon school and parental consent.

### Measures

We followed the same translation and back-translation procedures described in Study 1 to provide accurately worded items in Chinese to participants.

**Expressive flexibility.** Participants' abilities to flexibly enhance and suppress emotion expressions were measured with the newly developed CAFE Scale.

**Emotion expressiveness.** We adopted two scales to measure individual differences in habitual emotional expression: Emotion

Expressiveness Questionnaire (King & Emmons, 1990) and Berkeley Expression Scale (BES; Gross & John, 1995). Both are composed of 16 items and three dimensions. We used their two common dimensions in the current study, namely Positive Emotion Expressiveness and Negative Emotion Expressiveness. The third dimensions of these two scales are Intimacy and Impulse Strength, respectively, which were not included due to space eliminations and lack of obvious connection to the expressive flexibility construct. Before merging two scales, we first conducted an EFA with a total of 11 Positive Expressiveness items (e.g., "I often laugh so hard that my eyes water or my sides ache") and 10 Negative Expressiveness items (e.g., "Whenever I feel negative emotions, people can easily see exactly what I am feeling") from both scales. According to the results, there remained eight items and seven items under these two dimensions, respectively, all of which were rated from 1 = *not at all true of me* to 5 = *extremely true of me*. The Cronbach's alpha was .77 for Positive Emotion Expressiveness and .70 for Negative Emotion Expressiveness.

**Personality.** The short form of Big-Five Inventory-2 (BFI-2-S; Soto & John, 2017) was used to assess different aspects of personality, including Extraversion (e.g., "is outgoing, sociable"), Agreeableness (e.g., "is compassionate, has a soft heart"), Conscientiousness (e.g., "can sometimes be careless", reverse-scored), Neuroticism (e.g., "worries a lot"), and Openness to Experience (e.g., "is fascinated by art, music, or literature"). There are six items under each dimension, adding up to 30 items. Participants were asked to indicate to what extent each characteristic applied to them, ranging from 1 = *disagree strongly* to 5 = *agree strongly*. The Cronbach's alpha of the five subscales were .67, .55, .69, .76, and .75, respectively. The rather low reliabilities for three of these scales might be because early adolescents' self-descriptions develop in the midst of transitioning from a focus on concrete attributes toward more abstract characteristics (Montemayor & Eisen, 1977). The highly abstract items of this scale could potentially be difficult for them, compared with the more concrete behavior- and situation-based statements comprising other scales.

**Ego resilience.** We adopted the Ego-Resiliency Scale (Block & Kremen, 1996) to measure children's adaptability to changes in environment. Participants responded to 14 items (e.g., "I get over my anger at someone reasonably quickly") on a four-point scale, ranging from 1 = *does not apply at all* to 4 = *applies very strongly*. The Cronbach's alpha was 0.81 in the present study.

### Results

We first retested the factor structure and reliabilities of the CAFE Scale with the new sample. CFA was conducted again to examine the fitness of the dual-factor model. Results indicated an adequate fit, after adding three correlations between items within the same factor:  $\chi^2_{\text{SB}}(61) = 90.65$ , CFI = .94, TLI = .92, RMSEA = .04, SRMR = .06. The factor loadings of all items were above .40. The two subscales also showed adequate reliabilities, with Cronbach's  $\alpha = .74$  for Enhancement and .68 for Suppression.

We further assessed scale validity through relations between the CAFE scores and measures of Emotion Expressiveness, Personality, and Ego Resilience. As shown in Table 4, the CAFE Enhancement and Suppression subscales demonstrated some different correlation patterns. As expected, both Positive

Table 4  
*Measures of Validity: Means, Standard Deviations, and Correlations With CAFE Enhancement, Suppression, and Expressive Flexibility Scores (Study 2)*

Measure	<i>M</i>	<i>SD</i>	Bivariate correlations		
			Enhancement	Suppression	Expressive flexibility
Emotion expressiveness					
Positive emotion	3.56	.78	.15*	-.02	-.00
Negative emotion	2.78	.71	-.18**	-.21**	-.21**
Personality					
Extraversion	3.61	.74	.17**	.06	.11 <sup>†</sup>
Agreeableness	3.95	.56	.35***	.42***	.41***
Conscientiousness	3.56	.67	.27***	.44***	.44***
Neuroticism	2.52	.82	-.22***	-.44***	-.42***
Openness	3.70	.80	.26***	.24***	.25***
Ego resiliency	2.89	.51	.36***	.33***	.34***

Note. CAFE = Child and Adolescent Flexible Expressiveness Scale.

<sup>†</sup>  $p = .08$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Emotion Expressiveness and Extraversion only positively related to CAFE Enhancement score but showed nonsignificant correlations with CAFE Suppression. For the remaining dimensions of Emotion Expressiveness and Personality, CAFE scores demonstrated similar correlation profiles. Specifically, overall Expressive Flexibility and the two subcomponents were all negatively related with Negative Emotion Expressiveness and Neuroticism, whereas positively correlated with Agreeableness, Conscientiousness, and Openness to Experience. Finally, as expected, the Expressive Flexibility score showed a significant positive correlation with Ego Resilience, as did the two subscale scores.

## Discussion

Study 2 provided further support for the reliability and validity of the CAFE Scale using an urban sample, which is an important supplement to, and extension of, Study 1. The results of CFA were consistent with Study 1, showing that the factor structure was relatively stable. Although the internal consistency of suppression dimension was slightly below 0.70 ( $\alpha = .68$ ), generally the subscale reliabilities were acceptable.

The study further explored the convergent and discriminant validity of the CAFE Scale, which was not fully evidenced in Study 1. First, after distinguishing emotional expressivity by emotion valence, the differences between CAFE enhancement and suppression subscales were more obvious, particularly in terms of positive expressivity, which was only correlated with the enhancement score. Second, CAFE enhancement and suppression scores also showed different correlation patterns with the extraversion dimension of the Big Five. Again, extraversion was positively associated with enhancement, but showed a nonsignificant correlation with suppression. These two correlation patterns jointly provide further support to the notion that enhancement and suppression are two separate constructs. Third, overall expressive flexibility and the two subcomponents showed similar positive correlation patterns with adaptive indices, including agreeableness, conscientiousness, openness to experience, and ego resiliency, while also being negatively related to maladaptive indices such as neuroticism and negative emotion expressiveness. All these correlations were consistent with our expectations, which again sug-

gested the adaptive nature of expressive flexibility and the validity of the CAFE scale.

## Study 3: Comparing the CAFE Scale With the Expressive Flexibility Laboratory Task

Because the laboratory task developed by Bonanno's et al. (2004) is an objective and widely used paradigm to measure expressive flexibility, youth's observed performance in this task should be an important criterion for validating the new CAFE Scale. Previous research has demonstrated that the process of expressive regulation requires the involvement of conscious process (Richards & Gross, 2000). Therefore, individuals should be able to accurately assess their own abilities to regulate emotional expressions. Accordingly, we hypothesized that enhancement and suppression scores measured by the CAFE Scale would significantly predict corresponding performances in the laboratory task. Furthermore, the expressive flexibility scores from the scale and the task should be significantly related to each other.

## Participants and Procedure

The sample was from the first wave of the Facing Rejection Project, which is a broader longitudinal study focusing on Chinese children's emotions and interpersonal relationships. This project includes both primary and junior high school students, but only the latter group completed the CAFE Scale at Wave 1. Participants ( $N = 202$ ) of the present study were recruited from one rural junior high school and one urban junior high school of Shandong Province in China, on the condition of parental or guardian permission. None of the participants were involved in Study 1 or Study 2. We excluded data of three children because of obvious distraction during the task or failure to follow task instructions, so there remained 199 valid participants (48.2% female). Participants were in Grade 7 ( $n = 92$ ) or 8 ( $n = 107$ ), aged between 12 to 15 years ( $M_{\text{age}} = 13.57$  years,  $SD = 0.63$ ). They were first group-tested in quiet classrooms and completed a series of questionnaires, including the CAFE Scale. Within one week after the survey, children completed the laboratory task individually on a computer. Upon

completion, participants received small gifts (e.g., stationery sets or notebooks) for their participation.

## Measures

**Expressive flexibility scale.** We first used the CAFE Scale to measure children's ability to flexibly enhance and suppress emotional expressions. In this study, Cronbach's  $\alpha = .78$  for Enhancement and  $.70$  for Suppression.

**Expressive flexibility task.** The laboratory task established by Bonanno and colleagues (2004) was adopted. Participants viewed a series of emotion-eliciting pictures on a computer, while expressing emotions as requested. A camera placed just above the computer monitor filmed their facial expressions. To ensure that pictures used were suitable for children and could elicit strong emotions, we first conducted a pilot study for stimuli selection. There were in total six experimental blocks, with each block containing five pictures, so we eventually selected 15 positive pictures and 15 negative pictures for the formal expressive flexibility task that obtained the strongest respective valence ratings. All pictures were balanced across blocks according to the mean valence rated in the pilot study, such that all negative blocks and all positive blocks, respectively, were equivalent in their mean valence. We retained an additional set of five positive and five negative pictures as stimuli for two practice blocks. All picture stimuli and details about the picture selection are available from the authors on request.

We made several minor changes to adapt the task for a younger sample; we framed the task as a "game" that would be played with another child, provided extensive instructions about the meanings of the three different tasks, and began with practice trials to familiarize youth with the procedure. Before beginning, participants were told that they would complete an "emotion expression game" with another child of the same sex (not actually present) in the next room. Children would see blocks of pictures in sequence; the other child could not see the pictures, but he or she would sometimes see participants' facial expressions through the camera and attempt to guess their feelings on each block of pictures. One of three instructions would always appear ahead of each block, asking them to (1) enhance their expressions to let the observer easily guess their feelings, (2) suppress their expressions so that the observer could not guess their feelings, or (3) behave normally because the camera was switched off. Participants were told that they needed to try their best to behave in accordance with the instructions so that they could better cooperate with the observer to complete the game. After each block, participants rated their true feelings on a nine-point scale ( $-4 = \textit{extremely negative}$ ,  $0 = \textit{neutral}$ ,  $4 = \textit{extremely positive}$ ). To ensure that children could correctly follow the instructions, we added two practice blocks for enhancement and suppression tasks, respectively, followed by the formal task. Six blocks of stimuli randomly matched the three task instructions, with each instruction shown twice (once for positive stimuli and once for negative stimuli). The presentation time of each picture was 7 s, with a blank screen interval of 2 s between each photo.

Three trained coders with bachelor's or master's degrees in psychology rated participants' emotional expressivity for each block, ranging from 1 = *none* to 7 = *extreme*. Coders had never seen the emotional stimuli and were blind to the participants'

instructions. Agreement among the three coders was adequate, with an intraclass correlation coefficient of  $.92$ .

## Results

### Manipulation Check for the Laboratory Task

To ensure that the within-subject manipulation was effective, we respectively compared coder ratings and participants' subjective ratings across three tasks with repeated-measures analyses of variance (ANOVAs). Because the assumption of sphericity was violated in repeated-measures ANOVAs ( $ps < .001$ ), we reported results based on the Greenhouse-Geisser correction here. Results showed that (see Figure 2 in the online supplementary material), for coder ratings, expressivity scores significantly differed across three tasks,  $F(1.35, 267.50) = 256.53$ ,  $p < .001$ ,  $\eta^2 = .56$ . Further pairwise comparisons revealed that the mean scores of Enhancement ( $M = 4.20$ ), Normal ( $M = 3.87$ ), and Suppression ( $M = 2.55$ ) conditions were significantly different from each other (all  $ps < .001$ ). For subjective ratings (absolute values were used), the main effect of task was also significant,  $F(1.78, 352.09) = 9.90$ ,  $p < .001$ ,  $\eta^2 = .05$ . Participants' mean ratings in the Enhancement ( $M = 2.76$ ) and Normal conditions ( $M = 2.68$ ) were significantly higher than the Suppression condition ( $M = 2.49$ ,  $p < .001$  and  $p = .003$ ), but there was no significant difference between the former two ( $p = .120$ ). Although we expected nonsignificant differences for subjective ratings due to balanced valence between stimulus blocks, extensive research has also shown that modifying external emotion expressions can affect subjective responses to emotional stimuli (e.g., Adelman & Zajonc, 1989; Hawk, Fischer, & Van Kleef, 2012). In other words, participants may feel stronger subjective emotions when required to enhance expressions, while the intensity of subjective feelings would also decrease when required to suppress expressions. As shown from the means and effect sizes, however, the differences across three conditions of subjective ratings were much smaller than that of coder ratings. The overall pattern of results suggests that the manipulation was still effective.

### Comparing CAFE and Laboratory Task Scores

Because the laboratory task has been consistently used across most prior studies of expressive flexibility, it should be the criterion in this comparison and was therefore used as the dependent variable. Thus, we first performed two regression analyses that respectively used self-reported Enhancement and Suppression scores to predict expressivity in the Enhancement condition of the expressive flexibility task, while controlling for expressivity in the Normal condition of the task. This is also the analytical strategy utilized in the validation of the FREE Scale (Burton & Bonanno, 2016). Results showed that the CAFE Enhancement score positively predicted the mean expressivity rating in the Enhancement condition ( $\beta = .08$ ,  $p = .050$ ) after controlling the baseline expressivity level. In contrast, the CAFE Suppression score did not significantly predict expressivity in the Enhancement condition ( $\beta = .06$ ,  $p = .151$ ). We conducted similar regression analyses to predict expressivity scores in the Suppression condition. After controlling for expressivity in the Normal condition, higher Suppression scores measured by the CAFE Scale (indicating higher

suppression ability) significantly predicted lower expressivity in Suppression condition ( $\beta = -.14, p = .037$ ). Higher CAFE Enhancement scores also trended toward a negative association with expressivity in the Suppression condition ( $\beta = -.13, p = .054$ ). Finally, the correlations between Expressive Flexibility scores from the CAFE Scale and the laboratory task were calculated. Similar to the scale formula, we calculated the task flexibility score by subtracting the absolute value of the difference between Enhancement (indexed by differences on expressivity between Enhancement and Normal conditions) and Suppression scores (indexed by differences on Expressivity between Normal and Suppression conditions) from their sum. Results indicated that there was a significant positive correlation between the task and the scale Expressive Flexibility scores ( $r = .14, p = .048$ ).

### Discussion

Based on the first two studies, Study 3 attempted to establish the associations between the CAFE scores and children's observed abilities to enhance and suppress emotional expressions in a standardized laboratory task. This investigation is important for examining the convergent and discriminant validity of the newly developed scale, as the CAFE Scale was originally designed to be a convenient alternative to the laboratory task. Results indicated that expressive enhancement and suppression scores, as measured by the CAFE Scale, significantly predicted respective performances in the laboratory task. Specifically, after controlling for the baseline expressivity, self-reported enhancement scores positively predicted the degree of expressivity in the enhancement condition, and self-reported suppression scores negatively predicted the degree of expressivity in the suppression condition. The associations between CAFE scores and observed expressivity of the opposing conditions were not significant, although there was a trend toward an association between self-reported enhancement score and expressivity in the suppression condition. This is also reasonable, since expressive enhancement and suppression are still interrelated constructs. Higher self-reported suppression ability was always related to higher enhancement ability in our studies ( $r = .53, .53$ , and  $.50$ , from Study 1 to Study 3), so it is not surprising that they would share some associations with observed performance. Nevertheless, the two subscale scores of the CAFE Scale generally reflected the expected abilities and did not confound with each other. Importantly, the calculated flexibility scores from the scale and the laboratory task also showed a significant, positive association. Therefore, we may speculate that children are able to evaluate their own expressive flexibility, and the CAFE Scale could act as an eligible medium for such self-report.

It is also the case, however, that the strengths of links between the CAFE Scale and the laboratory task were more modest than those observed for the adult-reported FREE Scale. There are multiple differences in our scale construction and laboratory task procedures, compared with the validation of the FREE Scale, which might explain this finding. For example, the FREE Scale organizes enhancement and suppression items into separate sections of the measure and gives explicit instructions about which ability is concerned before each part, so individuals should have a very clear impression about the aim of each item. In contrast, the CAFE Scale randomly distributes enhancement and suppression items across the scale and incorporates the expected responses

with the scenarios to form final items. Although this format is likely easier for children to understand, it might also decrease the homogeneity between items and cause lower correlations with laboratory task performance. Psychologically, compared with that of adults, children's and adolescents' self-awareness is still undergoing development (Spencer, 1988). Thus, their self-perceptions might correspond more loosely to their actual abilities. Finally, in commenting on the fairly modest relationships between FREE Scale scores and laboratory task scores, Burton and Bonanno (2016) also noted that their self-report scale items were based on interpersonal contexts, while the laboratory task is highly asocial in nature. In this case, as well as for the CAFE Scale, it would be rather unsurprising for the self-report and laboratory scores to yield modest correlations. There seem to be both pros and cons for either measurement. The laboratory task might be more objective, but the inclusion of actual, youth-generated social scenarios in the self-reported scale improves its ecological validity. Both are likely to capture only a portion of youth's actual expressive flexibility in real-life contexts, suggesting a need for further research using a combined measurement approach.

### Study 4: Test-Retest Reliability

#### Participants, Measure, and Procedure

To assess test-retest reliability, a small fraction of participants in Study 1 were invited again 6 months later to complete the CAFE Scale. Because retesting all children who attended Study 1 was not feasible, we selected students in sixth grade as the target sample who were almost at the median age of all participants in Study 1. A total of 48 children (50.0% female) agreed to do the CAFE Scale again. This subset of participants aged between 12 and 13 years ( $M_{\text{age}} = 12.43$  years,  $SD = 0.38$ ). Similarly, they completed the CAFE Scale in a group, under the supervision of a trained instructor, and received another gift. The scale showed acceptable internal consistencies, with Cronbach's  $\alpha = .73$  for Enhancement and  $.68$  for Suppression.

#### Results and Discussion

We calculated correlation coefficients between participants' scores on the first and second administration of the CAFE Scale to determine the test-retest reliability. Results showed that the two Expressive Flexibility scores were strongly correlated with each other ( $r = .56, p < .001$ ). Significant correlations existed for the Enhancement subscale ( $r = .52, p < .001$ ) and for the Suppression subscale ( $r = .47, p = .001$ ). Although the value 0.70 is a common threshold for reliability assessment, this is not an absolute criterion. As Crocker and Algina (1986) suggested, it is difficult to set a fixed acceptable value for the test-retest reliability, since it is determined by a series of factors such as the interval time and sample type. In the present study, the period (6 months) between two measures was quite long and our participants were undergoing a developmental period with substantial changes in emotion sensitivity and self-regulation abilities, both of which might lead to lower test-retest correlations. In addition, the ability to regulate emotional expressions may not be fully crystallized. Even among adults tested across a 3-year period, expressive flexibility and its two subcomponents in the laboratory task showed only moderate

correlations ( $r = .40$  to  $.45$ ; Westphal et al., 2010). Considering such factors, we regarded these test–retest reliabilities in the current study to be acceptable.

### General Discussion

Expressive flexibility constitutes an important aspect of emotion regulation that is conceptually and behaviorally distinct from the regulation of subjective states (Zhu & Bonanno, 2017). The ease with which individuals can deploy different strategies to modify their overt emotional behavior, regardless of their internal feelings and cognitions, likely has important consequences for their everyday interpersonal interactions. Such abilities have also predicted psychosocial functioning across multiple studies using diverse samples. To date, however, investigations of this construct with children and adolescents have been rare. One potential reason might be the lack of appropriate measurement tools. The standard expressive flexibility task, which utilizes laboratory observation, is quite complicated, time-consuming, and rather artificial in nature. Moreover, the situations depicted in the existing self-report FREE Scale are not good fits with youth's experiences. Thus, a convenient and efficient measurement instrument targeting children and adolescents is an important step for exploring expressive flexibility and its clinical relevance in this group. To address this gap, we developed the CAFE Scale and provided support for its validity.

Considering ecological validity issues and the principle of conciseness, a scenario-based self-report scale would be a reasonable approach for examining youth's expressive flexibility. Therefore, we first conducted a pilot study to collect authentic scenarios in youth's lives that concerned expressive enhancement or suppression, which were then compiled into scale items. Study 1 identified seven enhancement items and six suppression items through factor analyses. Measurement invariance tests did not find significant differences between boys and girls, nor between primary and junior high school students. Both subscales showed adequate internal consistencies, and the CAFE scores showed expected relationships with several measures of psychosocial and behavioral adjustment, including depression, problem behaviors, and peer relationships. The majority of these correlations remained significant even after controlling for internal emotion regulation difficulties. Study 2 replicated the two-factor structure of the CAFE Scale in a new sample and demonstrated the differences between enhancement and suppression abilities by showing their differential links with positive emotion expressivity and extraversion. Study 3 further established the convergent and discriminant validity of the CAFE Scale by demonstrating the positive links between self-reported enhancement and suppression scores and observed emotional expressivity in corresponding conditions of the laboratory task but not the opposing conditions. Overall expressive flexibility scores, as measured by the scale and task, were also significantly correlated. This suggests that the CAFE Scale could be a potential alternative to the complex and time-consuming laboratory task. Finally, Study 4 demonstrated adequate test–retest reliability of the CAFE Scale by retesting a subsample of participants from Study 1.

Generally, across these various tests, the CAFE Scale showed good psychometric properties, including a stable factor structure, adequate internal consistencies, satisfactory test–retest reliability, as well as theoretically consistent convergent, discriminant, and

criterion-related validity. We found the hypothesized correlations between the CAFE and a series of relevant scales, as well as between children's CAFE scores and their actual performance in the laboratory task. In addition, all scale items were based on social scenarios generated by children, themselves, which suggests high ecological validity. Taken together, the CAFE Scale appears to be a reliable and valid measurement of children's and adolescents' expressive flexibility.

Nevertheless, it is important to note some limitations of the present study. First, within the framework of regulatory flexibility (Bonanno & Burton, 2013), the CAFE scale mainly measures children's ability to use expressive enhancement and suppression strategies (i.e., the repertoire component of the model), without considering context sensitivity and feedback components of the model. These latter two components are also vital to successful emotion regulation. Context sensitivity enables individuals to detect demands in ever-changing situations and choose the most appropriate regulatory strategy, whereas responses to social feedback help individuals to monitor and rapidly adjust their chosen strategies. In keeping with existing expressive flexibility measurements and the major focus of prior literature, the CAFE Scale focuses only on the repertoire component. Future research might aim to develop a more comprehensive instrument by considering these three components together, such as recording individuals' reactions in a relatively natural social context, and examining strategy choice, actual utilization, and subsequent adjustments. Alternatively, developing a measure specifically targeting context sensitivity or incorporation of feedback could allow researchers to make comparisons between these three components.

Second, the participants of all four studies were Chinese children and adolescents, which may limit the generalizability of the scale. Although this problem is unavoidable to some extent and exists any time a measure is developed and validated within a single cultural context, we also tried to minimize the cultural impact in several ways. First, we recruited participants from both rural and urban areas, to increase the diversity of our samples. Second, when selecting the scenarios, we considered the principle of universality; all the contexts depicted in the CAFE items are possible occurrences in the lives of children and adolescents from both Eastern and Western cultures. Although there might be cultural differences regarding the links between expressive flexibility and adjustment because of distinct cultural display rules for emotional expression (Matsumoto et al., 2008), this does not imply that the applicability of the scale would necessarily change across cultures. The FREE Scale is a good example, in that it was initially developed in a Western context but applied well to Chinese college students (Chen et al., 2018). Nevertheless, it is still necessary for future research to test the reliability and validity of the CAFE Scale in other cultures.

Third, as a self-reported scale, social desirability, insight deficits, or over-/underconfidence could influence children's responses. These issues may partly explain why younger participants showed higher CAFE scores in Study 1, though there are also potential developmental explanations for these age differences. These inherent drawbacks of self-reports might also be reflected in the modest correlations between CAFE scores and children's performances in the laboratory task. However, the introduction of the scale explicitly states that respondents should rate all items according to "whether you can do" instead of "whether you are

supposed to do,” which could remind participants to respond according to their true capabilities and decrease the influences of social desirability, to some extent.

Finally, the cross-sectional nature of all four studies limits conclusions about predictive validity. It would be valuable to investigate the associations between the CAFE scores and children’s psychosocial adjustment within a longitudinal design. This not only could further demonstrate the validity of the CAFE Scale, but also could address that, to date, there has been no direct empirical evidence showing the long-term adaptiveness of expressive flexibility for youth’s psychological adjustment. In addition, the developmental characteristics of expressive flexibility remain unclear, and a longitudinal study with several waves may provide relevant information.

Despite these limitations, the CAFE Scale appears to be a reliable and valid instrument for future research in the field of emotion regulation. Clinically, it could first be adopted independently to assess children’s and adolescents’ expressive flexibility. Additionally, considering its unique contributions to psychological functioning after controlling for internal regulation difficulties, researchers might also use the CAFE Scale in conjunction with other emotion regulation measures. Since existing scales often focus on internal emotion regulation, the CAFE Scale’s focus on external regulation is an important supplement that allows researchers to compare these two processes and to investigate children’s emotion regulation more comprehensively. In addition, the CAFE Scale might provide specific information regarding the emotional functioning of children and adolescents who show signs of poorer psychosocial adjustment. Identifying children with deficits in emotion expression and emotion regulation is an essential component of providing effective, targeted intervention to those who are experiencing difficulties.

## References

- Adelmann, P. K., & Zajonc, R. B. (1989). Facial efference and the experience of emotion. *Annual Review of Psychology*, *40*, 249–280. <http://dx.doi.org/10.1146/annurev.ps.40.020189.001341>
- Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review*, *30*, 217–237. <http://dx.doi.org/10.1016/j.cpr.2009.11.004>
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*, 411–423. <http://dx.doi.org/10.1037/0033-2909.103.3.411>
- Block, J., & Kremen, A. M. (1996). IQ and ego-resiliency: Conceptual and empirical connections and separateness. *Journal of Personality and Social Psychology*, *70*, 349–361. <http://dx.doi.org/10.1037/0022-3514.70.2.349>
- Bonanno, G. A., & Burton, C. L. (2013). Regulatory flexibility: An individual differences perspective on coping and emotion regulation. *Perspectives on Psychological Science*, *8*, 591–612. <http://dx.doi.org/10.1177/1745691613504116>
- 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*, 482–487. <http://dx.doi.org/10.1111/j.0956-7976.2004.00705.x>
- Bradburn, N. M. (1969). *The structure of psychological well-being*. Chicago, IL: Aldine.
- 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*, 929–941. <http://dx.doi.org/10.1037/pas0000231>
- Chaplin, T. M., Cole, P. M., & Zahn-Waxler, C. (2005). Parental socialization of emotion expression: Gender differences and relations to child adjustment. *Emotion*, *5*, 80–88. <http://dx.doi.org/10.1037/1528-3542.5.1.80>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*, 464–504. <http://dx.doi.org/10.1080/10705510701301834>
- 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. <http://dx.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*, 233–255. [http://dx.doi.org/10.1207/S15328007SEM0902\\_5](http://dx.doi.org/10.1207/S15328007SEM0902_5)
- Cillessen, A. H. N., & Bukowski, W. M. (2000). *Recent advances in the measurement of acceptance and rejection in the peer system. New directions in child and adolescent development, Number 88*. New York, NY: Jossey-Bass.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York, NY: Holt, Rinehart & Winston.
- Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature Reviews Neuroscience*, *13*, 636–650. <http://dx.doi.org/10.1038/nrn3313>
- Durán, J. I., Reisenzein, R., & Fernández-Dols, J. M. (2017). Coherence between emotions and facial expressions: A research synthesis. In J. M. Fernández-Dols & J. A. Russell (Eds.), *The science of facial expression* (pp. 107–129). New York, NY: Oxford University Press.
- Eisenberg, N., & Spinrad, T. L. (2004). Emotion-related regulation: Sharpening the definition. *Child Development*, *75*, 334–339. <http://dx.doi.org/10.1111/j.1467-8624.2004.00674.x>
- Farrell, A. D., Kung, E. M., White, K. S., & Valois, R. F. (2000). The structure of self-reported aggression, drug use, and delinquent behaviors during early adolescence. *Journal of Clinical Child Psychology*, *29*, 282–292. [http://dx.doi.org/10.1207/S15374424jccp2902\\_13](http://dx.doi.org/10.1207/S15374424jccp2902_13)
- Goldin, P. R., McRae, K., Ramel, W., & Gross, J. J. (2008). The neural bases of emotion regulation: Reappraisal and suppression of negative emotion. *Biological Psychiatry*, *63*, 577–586. <http://dx.doi.org/10.1016/j.biopsych.2007.05.031>
- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation: Development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*, *26*, 41–54. <http://dx.doi.org/10.1023/B:JOBA.0000007455.08539.94>
- Gross, J. J., & John, O. P. (1995). Facets of emotional expressivity: Three self-report factors and their correlates. *Personality and Individual Differences*, *19*, 555–568. [http://dx.doi.org/10.1016/0191-8869\(95\)00055-B](http://dx.doi.org/10.1016/0191-8869(95)00055-B)
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, *85*, 348–362. <http://dx.doi.org/10.1037/0022-3514.85.2.348>
- Gupta, S., & Bonanno, G. A. (2011). Complicated grief and deficits in emotional expressive flexibility. *Journal of Abnormal Psychology*, *120*, 635–643. <http://dx.doi.org/10.1037/a0023541>
- Haas, L. M., McArthur, B. A., Burke, T. A., Olino, T. M., Abramson, L. Y., & Alloy, L. B. (2019). Emotional clarity development and psychosocial outcomes during adolescence. *Emotion*, *19*, 563–572. <http://dx.doi.org/10.1037/emo0000452>
- Hawk, S. T., Fischer, A. H., & Van Kleef, G. A. (2012). Face the noise: Embodied responses to nonverbal vocalizations of discrete emotions. *Journal of Personality and Social Psychology*, *102*, 796–814. <http://dx.doi.org/10.1037/a0026234>

- Hawk, S. T., Keijsers, L., Branje, S. J. T., Graaff, J. V., Wied, M., & Meeus, W. (2013). Examining the Interpersonal Reactivity Index (IRI) among early and late adolescents and their mothers. *Journal of Personality Assessment*, *95*, 96–106. <http://dx.doi.org/10.1080/00223891.2012.696080>
- King, L. A., & Emmons, R. A. (1990). Conflict over emotional expression: Psychological and physical correlates. *Journal of Personality and Social Psychology*, *58*, 864–877. <http://dx.doi.org/10.1037/0022-3514.58.5.864>
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.
- Kovacs, M., & Beck, A. T. (1977). An empirical-clinical approach toward a definition of childhood depression. In J. G. Schulterbrandt & A. Askin (Eds.), *Depression in childhood: Diagnosis, treatment, and conceptual models* (pp. 1–27). New York, NY: Raven Press.
- Kring, A. M., Smith, D. A., & Neale, J. M. (1994). Individual differences in dispositional expressiveness: Development and validation of the Emotional Expressivity Scale. *Journal of Personality and Social Psychology*, *66*, 934–949. <http://dx.doi.org/10.1037/0022-3514.66.5.934>
- Le, B. M., & Impett, E. A. (2016). The costs of suppressing negative emotions and amplifying positive emotions during parental caregiving. *Personality and Social Psychology Bulletin*, *42*, 323–336. <http://dx.doi.org/10.1177/0146167216629122>
- Lev-Wiesel, R., Sarid, M., & Sternberg, R. (2013). Measuring social peer rejection during childhood: Development and validation. *Journal of Aggression, Maltreatment & Trauma*, *22*, 482–492. <http://dx.doi.org/10.1080/10926771.2013.785456>
- Matsumoto, D., Yoo, S. H., Nakagawa, S., & Multinational Study of Cultural Display Rules. (2008). Culture, emotion regulation, and adjustment. *Journal of Personality and Social Psychology*, *94*, 925–937. <http://dx.doi.org/10.1037/0022-3514.94.6.925>
- Montemayor, R., & Eisen, M. (1977). The development of self-conceptions from childhood to adolescence. *Developmental Psychology*, *13*, 314–319. <http://dx.doi.org/10.1037/0012-1649.13.4.314>
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Author.
- Richards, J. M., & Gross, J. J. (2000). Emotion regulation and memory: The cognitive costs of keeping one's cool. *Journal of Personality and Social Psychology*, *79*, 410–424. <http://dx.doi.org/10.1037/0022-3514.79.3.410>
- 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. <http://dx.doi.org/10.1016/j.jad.2016.09.027>
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Thousand Oaks, CA: SAGE.
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514. <http://dx.doi.org/10.1007/BF02296192>
- Schall, M., Martiny, S. E., Goetz, T., & Hall, N. C. (2016). Smiling on the inside: The social benefits of suppressing positive emotions in outperformance situations. *Personality and Social Psychology Bulletin*, *42*, 559–571. <http://dx.doi.org/10.1177/0146167216637843>
- Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*, 117–143. <http://dx.doi.org/10.1037/pspp0000096>
- Spencer, M. B. (1988). Self-concept development. *New Directions for Child and Adolescent Development*, *42*, 59–72. <http://dx.doi.org/10.1002/cd.23219884206>
- Thompson, R. A. (1994). Emotion regulation: A theme in search of definition. *Monographs of the Society for Research in Child Development*, *59*(2–3), 25–52. <http://dx.doi.org/10.1111/j.1540-5834.1994.tb01276.x>
- Trommsdorff, G., & Rothbaum, F. (2009). Development of emotion regulation in cultural context. In M. Vandekerckhove, C. Von Scheve, S. Ismer, S. Jung, & S. Kronast (Eds.), *Regulating emotions: Culture, social necessity, and biological inheritance* (pp. 85–120). Cambridge, MA: Blackwell Publishing.
- Westphal, M., Seivert, N. H., & Bonanno, G. A. (2010). Expressive flexibility. *Emotion*, *10*, 92–100. <http://dx.doi.org/10.1037/a0018420>
- Wong, E., Tschann, F., Messerli, L., & Semmer, N. K. (2013). Expressing and amplifying positive emotions facilitate goal attainment in workplace interactions. *Frontiers in Psychology*. Advance online publication. <http://dx.doi.org/10.3389/fpsyg.2013.00188>
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, *34*, 806–838. <http://dx.doi.org/10.1177/0011000006288127>
- Zhu, Z., & Bonanno, G. A. (2017). Affective flexibility: Relations to expressive flexibility, feedback, and depression. *Clinical Psychological Science*, *5*, 930–942. <http://dx.doi.org/10.1177/2167702617717337>
- Zimmermann, P., & Iwanski, A. (2014). Emotion regulation from early adolescence to emerging adulthood and middle adulthood: Age differences, gender differences, and emotion-specific developmental variations. *International Journal of Behavioral Development*, *38*, 182–194. <http://dx.doi.org/10.1177/0165025413515405>

Received February 19, 2019

Revision received August 11, 2019

Accepted November 20, 2019 ■