Emotion
2011, Vol. 11, No. 2, 215–223

1528-3542/11/$12.00 DOI: 10.1037/a0022602

Measuring Children's Regulation of Emotion-Expressive Behavior

Yair Bar-Haim, Gali Bar-Av, and Avi Sadeh
Tel Aviv University

Emotion regulation has become a pivotal concept in developmental and clinical research. However, the measurement of regulatory processes has proved extremely difficult, particularly in the context of within-subject designs. Here, we describe a formal conceptualization and a new experimental procedure, the Balloons Game, to measure a regulatory component of emotion-expressive behavior. We present the internal consistency and stability of the indices derived from the Balloons Game in a sample of 121 kindergarten children. External validation against measures that have been associated with emotion regulation processes is also provided. The findings suggest that the Balloons Game provides a reliable tool for the study of regulation of emotion expression in young children.

Keywords: emotion regulation, children, reactivity, measurement

Over the past two decades, the concept of emotion regulation has received an increasing degree of theoretical and empirical consideration (e.g., Cole, Martin, & Dennis, 2004; Fox, 1994; Gross, 2002; Gross & Levenson, 1993; Thompson, 1994) and has become a pivotal organizing concept in developmental and clinical research. Although many researchers readily agree that emotion regulation is of central importance, assessing the extent of emotion regulation in any particular instance has proved extremely elusive. Here, we propose a new procedure to measure a regulatory component of emotion-expressive behavior. Specifically, we focus on the formal estimation of effortful suppression of emotion-expressive behavior as an index of emotion regulation. This form of emotion regulation has been defined as the conscious inhibition of emotional expressive behavior while emotionally aroused (Gross & Levenson, 1993).

Our selection of emotion-expressive behavior as an index of regulatory function was guided by the observation that people are often required to actively inhibit their ongoing emotion-expressive behavior, and their ability to do so has been noted by many as an important developmental milestone (e.g., Calkins & Fox, 2002; Kopp, 1989; Thompson, 1991). In addition, under- or overregulation of emotion-expressive behavior are considered key diagnostic features in a wide array of psychiatric disorders associated with emotional instability, such as mood and anxiety disorders (Barlow, 1991; Gross, 1998; Gross & Munoz, 1995).

A major challenge to the study of emotion regulation has been formulated as the difficulty in distinguishing emotion regulation from emotional reactivity (e.g., Davidson, Jackson, & Kalin, 2000; Kagan, 1994; Stansbury & Gunnar, 1994). In the context of emotion-expressive behavior, this formulation implies that any observed behavior must involve at least two processes: the regulated emotional reaction and the regulatory process itself (e.g., Cole et al., 2004; Derryberry & Rothbart, 1997; Frijda, 1988; Gross, 1999). These two processes are not readily separable in observed behavior. Partly, this difficulty reflects a lack of consensus regarding the nature of emotion and partly an inability to stipulate the specific nature of the emotion-regulation process. Of particular concern is the fact that a large portion of the studies using the term emotion regulation fail to supply evidence for the existence of a regulatory process and do not provide external validation to their measures (Cole et al., 2004).

Evidence of an emotion regulatory process may be obtained by showing a change in an outcome measure of emotion expression under two distinct levels of experimentally manipulated emotion regulation. Specifically, it may be assumed that individuals regulate emotions even when they are not instructed to do so, and that the relative increase in emotion regulation during a period when they are instructed to regulate, compared to a baseline of spontaneous emotion regulation, may be operationalized as being propositional to emotion regulation capacity. Only a handful of studies have manipulated the level of required emotion regulation while attempting to keep constant the assumed emotional reactivity (e.g., Cole, 1986, Exp. 2; Gross & Levenson, 1993; Richards & Gross, 2000). However, these studies did not estimate the unique contribution of emotion regulation to the observed emotion-expressive behavior within subjects and, instead, relied on between-groups designs. Although between-groups research designs have a number of advantages in the study of emotion regulation (e.g., reducing demand effects), within-subject measures are necessary for practical, individually based applications. The scarcity of within-subject designs on emotion regulation may be attributed, in part, to the fact that many emotion-provoking procedures cannot be used repetitively either because once the manipulation has been applied it might not provoke the same emotion again, or it might not be ethically applied more than once (e.g., the disappointment paradigm; Saarni, 1984).

Yair Bar-Haim, Gali Bar-Av, and Avi Sadeh, The Adler Center for Research in Child Development and Psychopathology, Department of Psychology, Tel Aviv University, Tel Aviv, Israel.

Correspondence concerning this article should be addressed to Yair Bar-Haim, The Adler Center for Research in Child Development and Psychopathology, Department of Psychology, Tel Aviv University, P.O. Box 39040, Tel Aviv 69978, Israel. E-mail: yair1@post.tau.ac.il

Below, we describe a new procedure for assessing the regulation of emotion-expressive behavior in children. In our view, this procedure overcomes many of the technical difficulties in the measurement of regulation of emotion-expressive behavior in children to date. Specifically, we designed a task and a measurement system that (a) repeatedly elicits both positive and negative emotions, (b) allows repeated measurements of emotion-expressive behavior, (c) tailors task difficulty individually, thereby controlling for individual variability in performance, (d) experimentally manipulates two distinct levels of emotion-expression regulation, and (e) generates multiple emotion regulation indices that take into account both within individual changes in regulation upon demand, as well as the child's general levels of emotion expressivity relative to other children. Importantly, however, the theoretical framework and the experimental solution offered here should not be taken as an attempt to resolve the emotion regulation riddle by isolating emotion regulation from emotional reactivity, but rather as a proof of concept that a measure that is proportional to emotion regulation can be derived. Before we describe this method in greater detail, we demonstrate how an estimation that is proportional to emotion regulation may be formally derived.

Formal Estimation of Emotion Regulation

Consider, as an example, two experimental conditions designed to elicit frustration and anger: In the low-regulation condition, frustration and anger are experimentally elicited, but no particular instruction concerning expressive behavior is provided. By contrast, in the high-regulation condition, the child is frustrated and made angry in the same way, but is explicitly instructed not to reveal his or her emotions. We assume that the observed behavior (OB) of the child after being frustrated and made angry is the sum of his or her emotional reactivity (E) and an emotion regulatory process (R) he or she employs. Specifically, R represents the percent of behavioral suppression out of the maximal unregulated expressive behavior. Thus, the value of R can increase when greater suppression is required, but cannot exceed 100% of the unregulated behavior. It may be further assumed that the recruited regulatory process is modulated by the intensity of the emotional reaction.

In the low emotion regulation condition, this may be expressed as Observed Behavior = Emotional Reactivity – Emotion Regulation × Emotional Reactivity, where the values of R range between 0 and 1, or as shown in Equation 1:

$$OB_{\text{low regulation}} = E - (R * E) \tag{1}$$

In the high emotion regulation condition, we assume the same emotional reactivity (E), but a higher degree of employed emotion regulation, nR, where $n > 1$ and indexes the additional amount of successful emotion suppression (i.e., n increases with increased experimental demand for regulation). This may be expressed as shown in Equation 2:

$$OB_{\text{high regulation}} = E - (nR * E) \tag{2}$$

Therefore, a measure that is proportional to emotion regulation may be empirically estimated by dividing Equation 1 by Equation 2. This may be expressed as shown in Equation 3:

$$\begin{aligned} OB_{\text{low regulation}}/OB_{\text{high regulation}} &= E\,(1 - R)/E\,(1 - nR) \\ &= \mathbf{(1 - R)/(1 - nR)} \end{aligned} \tag{3}$$

Goals and Hypotheses

With the above formal derivations at our disposal, the aim of this study was to establish the validity of a new procedure, the *Balloons Game* (see Method for a full description), both as a measure of emotion expressivity and, more important, as a measure of emotion regulation in children. To that end, we tested four primary hypotheses. The first three hypotheses were related to the underlying assumptions of our conceptual model, and the fourth was focused on external validation of the emotion regulation indices derived from the Balloons Game.

First, we expected children to display considerable trial-by-trial stability in emotion-expressive behavior in the relatively standard context of the task. Second, we expected that children's ability to regulate positive and negative emotional expressivity under conditions of both on-task success and failure would correlate significantly, reflecting a core regulatory mechanism. Third, we expected that, as task requirements for regulation increased, correlations between our emotion regulation indices and raw emotional expressivity would also increase. Specifically, if observed behavior (i.e., raw expressivity scores) reflects emotional reactivity to the task minus a regulatory effort, then an experimentally induced increase in regulatory effort should increase the common variance between the total observed behavior (i.e., raw expressivity) and the regulatory process (i.e., emotion regulation).

Finally, we expected our derived estimates of emotion regulation to correlate with other measures that have been implicated conceptually and empirically in emotion regulatory capacity. Such external validation for emotion regulation measures is scarce in the extant literature. For the present study, we selected three types of external measures: parental reports on children's temperamental tendencies (effortful control, attentional focusing, and activity level) and on externalizing behavior problems, regulation of emotion during a mother-child interaction, and heart period.

Children are expected to regulate their emotion expressive behavior in various contexts. Difficulties in emotion regulation can manifest as externalizing behavior and high activity level (e.g., Eisenberg et al., 2001; Rubin, Coplan, Fox, & Calkins, 1995), whereas good attentional control and high effortful control have both been suggested as important components of good emotion regulation (Posner & Rothbart, 2000). We, therefore, expected our behaviorally derived indices of emotion regulation to negatively correlate with parental reports of children's externalizing behavior problems (Achenbach, 1991) and their temperamental activity level (Rothbart, Ahadi, Hershey, & Fisher, 2001) and positively correlate with good capacities for inhibitory control and attentional focusing (Rothbart et al., 2001).

We also expected our emotion regulation indices to positively correlate with the ratings of the emotional tone of mother-child interaction in a task that challenged the dyad's cooperation skills and required high levels of emotion regulation from both participants (Falender & Heber, 1975). Because it is believed that parents provide young children with considerable emotion regulatory support, we hypothesized that actively challenging parent–child cooperation would form a good context to measure a correlate of emotion regulation (Allhusen et al., 2004; Calkins & Johnson, 1998; Oppenheim, Nir, Warren, & Emde, 1997; Shipman & Zeman, 2001).

Last, we expected to find associations between our emotion regulation indices and children's baseline pattern of cardiac activity (heart period). Heart period has been associated with various aspects of behavioral and emotion regulation and, specifically, with emotion-expressive behavior (e.g., Calkins, Graziano, Berdan, Keane, & Degnan, 2008; Cole, ZahnWaxler, Fox, Usher, & Welsh, 1996; Giuliani, McRae, & Gross, 2008; Gross & Levenson, 1993).

Method

Participants

A total of 121 children (58 males, 63 females) and their mothers participated in the study. Data were collected during the summer preceding the children's entrance to first grade. Children's mean age = 6.20 years, *SD* = 0.30, range = 5.50–6.80, mothers' mean age = 38.03, *SD* = 4.80. The children were recruited from mainstream public kindergartens in 11 different cities in the greater Tel Aviv area, Israel. Letters explaining the study were delivered to the parents, and those who agreed to participate by returning a signed informed consent by mail were recruited. The study was approved by the institutional ethics committee and the Israeli Ministry of Education Review Board.

Measures

The balloons game. The Balloons Game[1] is a simple computer-based task in which children are instructed to "pop" all green balloons by clicking on them with the computer's mouse. The green balloons are interspersed with balloons of different colors, all floating up the screen from the bottom and disappearing at the top (see Figure 1). The procedure begins with a 90-s calibration phase, during which a personally tailored game-speed level is determined for each child. The speed is tailored to be fast enough to challenge and engage the player, but slow enough to ensure success. Specifically, the successful "popping" of three green balloons increases game speed and three misses decrease it. At the end of the 90-s calibration period, a constant that is proportional to the child's performance was added to the final performance speed to further slow the determined game speed, thereby affording the child a near-perfect success rate. With this method, baseline differences in children's playing competence were controlled.

Figure 1. The Balloons Game. Children are instructed to pop all the green balloons to win the game and become the balloons champion. The red timer bar at the bottom of the screen indicates progression on the trial.

During the game proper, children were again instructed to pop all the green balloons appearing on the screen during a 40-s time period delineated by a visible red timer bar on the bottom of the screen. Success trials ended with rewarding music and a happy picture and were designed to invoke joy and pride. Interspersed between success trials, frustration and anger were induced when, toward the end (final 10 seconds) of an otherwise successfully proceeding trial, the child's performance was thwarted by a rigged failure of the computer mouse, leading to a failed trial. Failed trials were accompanied by a failure sound and picture designed to evoke frustration and anger. This computer game is highly engaging for children, as it evokes easily detected positive and negative emotion-expressive behavior, and it allows for repeated trials, thereby affording within-subject measurement of behavior.

To manipulate the level of required emotion regulation, each child was tested under two experimental conditions (low regulation and high regulation), the order of which was counterbalanced across participants. In the low-regulation condition, children played the Balloons Game, and no specific instructions were given. In the high-regulation condition, children were explicitly asked to suppress any display of emotions, so that "no one would be able to know by their behavior whether they are winning or losing the game." Each session consisted of eight 40-s trials. The first two trials in a session were always success trials. The remaining six trials consisted of three success trials and three failure trials presented in a pseudorandom order (e.g., SS-SFFSFS). The entire procedure was videotaped, and a frame-by-frame coding of children's facial expressions was applied, using the Facial Action Coding System (FACS; Ekman, Friesen, & Hager, 2002). Measures proportional to emotion regulation were computed by dividing observed behavior (i.e., specific FACS action units) under the low-regulation condition ($OB_{low\ regulation}$) by observed behavior (the same FACS action units) under the high-regulation condition ($OB_{high\ regulation}$).

FACS coding and derivation of expressivity and emotion regulation indices. Children's facial expressions were coded on a presence/absence basis, using a selected set of 14 action units (AUs) from the FACS (Ekman et al., 2002). Ten action units (AU4, AU5, AU7, AU9, AU10, AU15, AU16, AU17, AU23, and AU24) were selected to index negative emotional displays. Four action units (AU1, AU2, AU6, and AU12) were selected to index positive emotional display. These specific action units were selected based on reported associations between these and specific negative or positive emotional displays (Ekman et al., 2002) and

[1] Hebrew and English versions of the protocol and software for running the Balloons Game may be obtained from the authors.

based on previous work on emotion regulation with children (Cole, 1986; Cole, Teti, & Zahn-Waxler, 2003). For each condition, (i.e., high/low regulation), six trials of the balloons game were coded (three success trials and three failure trials). For every trial, we coded the 6 s preceding the sound and picture denoting success or failure and the 6 s after that sound and picture. This 12-s epoch captured the 6 s during which children were actively dealing with the induced mouse failure (fail trials) or the final seconds before formal success (success trials) and the 6 s after actual failure/success on a given trial. Within each second, five picture frames were coded (in intervals of 20 ms). Therefore, for each of the two regulation conditions, we obtained data from 60 coding frames from every trial of the Balloons Game (5 frames × 12 s), with a total of 180 coding units for the success trials and 180 coding units for the failure trials.

Derivation of emotion regulation indices. FACS coding was carried out by two graduate psychology students who reached coding reliability with a third, highly experienced coder. The two coders were blind to the aims of the study and unaware of the condition (high/low regulation) they were coding in any given tape. Care was taken so that a coder would code the videotape of a particular child in only one of the conditions (either low or high regulation). To assess coding reliability, 20% of the tapes were double-coded by one of the two graduate coders and the more experienced coder. Mean intercoder kappa scores across the entire matrix was 0.92 (range = 0.85–1.00) for one coder and 0.93 (range = 0.86–1.00) for the other coder. Discrepancies in coding were discussed and resolved by consensus.

Summing the occurrence frequencies of the different AUs across positive and negative expressions, failure and success conditions, and high- and low-regulation demands yielded eight raw scores: negative expressions during success, negative expressions during failure, positive expressions during success, and negative expressions during success for the high- and low-regulation conditions.[2] Five indices were computed for the low-regulation condition and five for high-regulation condition: Negative Expressivity (sum of negative expressions under failure and negative expressions under success); Positive Expressivity (sum of positive expressions under failure and positive expressions under success); Expressivity under Failure (sum of negative expressions under failure and positive expressions under failure); Expressivity under Success (sum of negative expressions under success and positive expressions under success); and Total Expressivity (sum of negative expressions under failure and success and positive expressions under failure and success).

In addition, five indices of regulation of emotion-expressive behavior (from now on, termed indices of emotion regulation) were derived by dividing observed behavior under high regulation by observed behavior under low regulation (see Introduction for a formal derivation). Specifically, Emotion Regulation Negative (Negative Expressivity high regulation/Negative Expressivity low regulation), Emotion Regulation Positive (Positive Expressivity high regulation/Positive Expressivity low regulation), Emotion Regulation Failure (Expressivity under Failure high regulation/Expressivity under Failure low regulation), Emotion Regulation Success (Expressivity under Success high regulation/Expressivity under Success low regulation), and Emotion Regulation Total (Total Expressivity high regulation/Total Expressivity low). Each of these five indices was multiplied by -1 *to create scales in which higher scores represent greater emotion regulation*.

Emotional tone of mother-child interaction. To assess the emotional tone of mother-child interaction in a context requiring emotion regulation and high levels of cooperation, we gave mother-child dyads a joint assignment (see Falender & Heber, 1975). Dyads were introduced to an etch-a-sketch magnetic drawing board with two knobs, one controlling horizontal cursor movement and one controlling vertical cursor movement. To draw oblique lines, both knobs must be rotated together. Mother and child were instructed to copy four geometric figures presented to them one by one. Two simple figures were introduced first and were used as practice trials, followed by two more complex figures. The mother was assigned to manipulate one of the knobs, and the child was assigned to manipulate the other. The dyads were allowed to discuss and plan their moves, but the mother and child were asked to only touch their assigned knob, therefore requiring high levels of cooperation and communication. The entire session was videotaped.

The emotional tone of dyadic interaction was coded on a 5-point scale, ranging from 1 (Highly Negative/Emotionally Dysregulated Interaction; e.g., shouting at each other, one dyad member controls both knobs, angry and upset facial expressions, pushing and shoving, degrading remarks, impatience), to 5 (Highly Positive Emotionally Regulated Interaction; e.g., mutual compliments, laughs, applauds, constructive planning and discussion). The index of mother-child interaction emotional tone was taken as the mean score for the two more complex figures copied. All dyads were able to complete the task, albeit with differing speeds and accuracies. As we were interested in the emotional tone of the interaction, we do not report measures of efficacy and performance. Videotapes were coded by two undergraduate psychology students. Twenty-five percent of the tapes were double-coded for reliability. Percent agreement ranged from 94% to 100%. Discrepancies in coding were discussed and resolved by consensus.

Heart-period (HP) acquisition. Electrocardiography (ECG) data were collected during a 5-min resting baseline condition. Two electrodes were placed on opposite sides of the child's torso, with the ground electrode being placed on the child's neck. ECG data were collected using the Ceegraph EEG system from Bio-Logic Systems Corp. (Mundelein, IL). The sampling rate was 256 Hz, with 16-bit resolution.

R-wave detection (to the nearest ms) was carried out using IBI analysis software from Orgil (Ein-Ayala, Israel). This program provides a data file containing the onset times of each detected R-wave in the record. For artifact editing, the sampled ECG signal was viewed graphically alongside tick marks representing the times of software-detected R-waves. In the case of an obscured R-wave that was not detected by the software, a tick mark was

[2] From these raw scores, various indices of emotion regulation and of raw expressive behavior may be computed, depending on the particular research aims of the user. As may be expected, the more inclusive the index, the greater its stability. Because the main aim of the present study was to demonstrate the diversified potential of the Balloons Game for research, we computed 10 somewhat overlapping aggregate indices of *raw emotion expressivity*, covering the main effects of emotion valence (negative, positive) and task manipulation (failure, success), as well as a total overall expressivity score.

inserted into the graphical ECG record. If the undetected R-wave was visible, it was marked manually; if the R-wave was not visible, the tick mark was placed based on specific editing rules (Byrne & Porges, 1993). For each child, mean HP was calculated for the 5-min recording.

Parental Reports

Child Behavior Checklist (CBCL). Children's behavior problems were assessed using the Hebrew version of the CBCL. The CBCL (Achenbach, 1991) is a 118-item parent-report measure of child behavior problems. The checklist provides scores on several behavior problem areas and identifies broadband indices of internalizing and externalizing behavior problems. The broadband indices were used for subsequent analyses. The Hebrew version of the CBCL was found to be valid, internally consistent (Zilber, Auerbach, & Lerner, 1994), and reliable in distinguishing between patient and control groups (Apter, Offer, Blumenson, & Tyano, 1988). Cronbach's alphas for the internalizing and externalizing scales for the present sample were 0.87 and 0.86, respectively.

Child Behavior Questionnaire (CBQ). Children's temperament was assessed using the CBQ (Rothbart et al., 2001). The CBQ includes 80 items that yield eight scales pertaining to different dimensions of child temperament. We were interested in three scales of the CBQ that index temperamental features, which have been theoretically associated with the following emotion regulation capacities: attentional focusing: capacity to maintain attentional focus on task-related channels; inhibitory control: capacity to plan and suppress inappropriate approach responses under instructions or in novel or uncertain situations; and activity level: gross motor activity, including rate and extent of locomotion (Derryberry & Rothbart, 1997; Posner & Rothbart, 2000). Cronbach's alphas for the attention-focusing, inhibitory-control, and activity-level scales for the present sample were 0.81, 0.80, and 0.86, respectively.

General Procedure

After informed consent was obtained, children and their mothers were invited to visit the laboratory at the university. Each visit began with the Balloons Game, administered in a quiet testing room by an experimenter who was quietly present throughout the procedure. Mothers were escorted to an adjacent room and filled out the questionnaires packet. After the Balloons Game, the children were allowed a break and a rest to assure that they had recovered from the potential frustration experienced during the Balloons Game. This break was followed by a 5-min recording of resting baseline ECG. For ECG recording, children sat quietly on a comfortable reclining chair and were asked to avoid moving. To assure that all participants were in a quiet, awake state, they were shown a short animated movie during ECG collection and were told to pay attention, as, later, they would be asked a few questions about the film's content. Finally, quality of mother-child interaction was assessed. Mother and child were seated in front of a small table situated in a play room, and were asked to copy geometric shapes using the coordinated etch-a-sketch procedure described above. A video camera was placed in front of the dyad, adjusted to record a full picture of the situation. An experimenter was present in the room throughout the procedure and handed out the figures to be copied, but did not interfere with the action.

Results

Preliminary Analyses

To determine whether task presentation order (i.e., high- or low-regulation condition first) had an effect on children's emotion-expressive behavior, we contrasted the raw expressivity and emotion regulation scores between these two counterbalanced conditions. The results indicate nonsignificant order effects, all $ps > .35$.

We contrasted the raw negative, positive, failure, success, and total emotion expressivity scores in the low- and high-regulation conditions to test children's responsiveness to our regulatory request (see Figure 2). Children responded with significant decreases in emotion-expressive behavior to our experimental instruction to suppress emotional expressivity, all $ps < .0001$.

Finally, we tested for gender differences by contrasting boys' and girls' emotion regulation scores, as well as their raw emotion-expressivity scores, in the low- and high-regulation conditions. These contrasts revealed that boys and girls did not differ on any of the emotion regulation indices, all $ts(119) < 1.60$, $ps > .10$, except for raw expressivity under success conditions. When successful on the Balloons Game, boys were more emotionally expressive than girls: In the low-regulation condition ($M = 112.53$, $SD = 61.69$ vs. $M = 90.57$, $SD = 49.79$, respectively), $t(119) = 2.16$, $p < .05$; in the high-regulation condition ($M = 78.57$, $SD = 52.40$ vs. $M = 56.89$, $SD = 43.78$, respectively), $t(119) = 2.48$, $p < .01$.

Stability of Raw Expressivity Scores Across Experimental Trials

We expected children to display considerable trial-by-trial stability in emotion-expressive behavior. An inspection of Table 1 reveals significant stability in raw emotion-expressive behavior, confirming that these behaviors are quite stable when measured repeatedly in conditions with identical emotional valance (e.g., success low regulation correlated with success high regulation, or failure low regulation correlated with failure high regulation). Considerable associations, but of smaller magnitude, were also detected across emotional valences, suggesting that the general tone of emotion expressivity is, at least partially, governed by a common mechanism that is valance- and context-independent. In

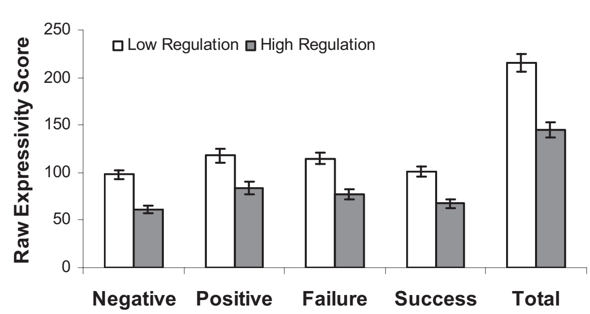

Figure 2. Mean raw expressivity scores (negative, positive, failure, success, and total) and standard error bars in the low and high emotion regulation conditions. All the differences between the low- and high-regulation conditions were significant at $p < .0001$.

Table 1
Correlations Between Raw Negative Expressivity Scores (Top Section) and Raw Positive Expressivity Scores (Bottom Section) on Repeated Failure and Success Trials, During Low or High Emotion-Regulation Demands

	Fail			Success		
	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3
Negative expressions						
Low regulation						
Fail						
Trial 1	—					
Trial 2	$.67^{***}$	—				
Trial 3	$.46^{***}$	$.62^{***}$	—			
Success						
Trial 1	$.25^{**}$	$.19^{*}$	.11	—		
Trial 2	.14	$.22^{*}$	.13	$.29^{***}$	—	
Trial 3	$.18^{*}$	.12	$.23^{*}$	$.29^{***}$	$.42^{***}$	—
High regulation						
Fail						
Trial 1	—					
Trial 2	$.59^{***}$	—				
Trial 3	$.59^{***}$	$.80^{***}$	—			
Success						
Trial 1	.01	$.24^{**}$	.14	—		
Trial 2	.09	.08	.09	$.54^{***}$	—	
Trial 3	.01	.17	.12	$.62^{***}$	$.51^{***}$	—
Positive expressions						
Low regulation						
Fail						
Trial 1	—					
Trial 2	$.62^{***}$	—				
Trial 3	$.78^{***}$	$.72^{***}$	—			
Success						
Trial 1	$.25^{**}$	$.29^{***}$	$.30^{***}$	—		
Trial 2	$.18^{*}$	$.25^{**}$	$.31^{***}$	$.63^{***}$	—	
Trial 3	$.35^{**}$	$.35^{***}$	$.34^{***}$	$.50^{***}$	$.62^{***}$	—
High regulation						
Fail						
Trial 1	—					
Trial 2	$.81^{***}$	—				
Trial 3	$.71^{***}$	$.74^{***}$	—			
Success						
Trial 1	$.30^{***}$	$.35^{***}$	$.32^{***}$	—		
Trial 2	$.28^{**}$	$.31^{***}$	$.31^{***}$	$.62^{***}$	—	
Trial 3	$.20^{*}$	$.24^{**}$	$.28^{**}$	$.53^{***}$	$.73^{***}$	—

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

conclusion, strong stability of raw expressive behavior may be inferred from the data.

Associations Among Indices of Regulation of Emotion-Expressive Behavior

An inspection of Table 2 reveals highly significant correlations between the different indices of emotion regulation across conditions, *r*s range between 0.49 and 0.84, all *p*s $<$ .0001. This correlations pattern possibly reflects the operation of a core regulatory capacity across conditions. A notable exception is the lower, trend level correlation between emotion regulation of negative and positive facial expressions. Possible implications of this particular association are discussed below.

Associations Among Raw Expressivity Scores and Indices of Emotion Regulation

Additional support for the validity of our theoretical formulation and experimental manipulation could be provided if correlations between raw expressivity scores and emotion regulation indices would be stronger under the high-regulation than the low-regulation condition.[3] That is, if, at every instance of observed behavior (i.e., raw expressivity scores), we, in fact, were to observe emotional reactivity to the task minus a regulatory process, then an experimentally induced increase in regulatory process (i.e., our instruction to suppress expressive behavior) should increase the common variance between the total observed behavior (raw expressivity indices) and the regulatory process (emotion regulation indices). An inspection of Table 3 reveals nonsignificant or sporadic correlations of low magnitude between raw emotion expressivity scores under low regulatory demands and the emotion regulation indices. By contrast, highly significant correlations emerge between all the indices of emotion regulation and raw expressivity scores in all conditions under high emotion regulatory demands. Fisher's *r*-to-*Z* comparisons indicate significant differences in the magnitudes of these two classes of correlations for congruent conditions (all *p*s $<$ 0.001).

External Validation of Emotion Regulation Indices Derived From the Balloons Game

We expected our newly derived estimates of emotion regulation to correlate with other measures that have been previously associated with different aspects of emotion regulatory capacity in children. The data (see Table 4) show that the emotion regulation indices correlate significantly with almost all the preselected external measures of emotion regulation.

Although some of the subindices of emotion regulation correlated significantly with specific external measures (e.g., ER of positive expressions and ER during success trials with CBQ–Inhibitory control), whereas others did not (e.g., ER of negative expressions and ER during failure trials), the correlation pattern was the same for all ER measures, and *r*-to-*Z* comparisons indicated no significant difference in the magnitude of these correlations. Thus, focusing on the total emotion regulation index (the most stable, inclusive of the ER indices), suggests that high emotion regulatory capacity (a high score on the index) is associated with lower levels of externalizing behavior problems, lower activity levels, and higher capacities for inhibitory control and attentional focusing, as reported by the parents. In addition, higher emotion regulatory capacity is correlated with higher quality (i.e., better emotional tone) of mother–child interaction during a demanding cooperative task and with shorter heart period (i.e., faster heart rate).

[3] Note that emotion regulation scores and raw expressivity scores are not independent variables, and this calls for caution in results interpretation. However, the presented analyses provide an estimate of the extent of this dependency between raw expressivity and regulation under different regulatory requirements.

Table 2
Means and Standard Deviations of the Emotion Regulation (ER) Indices Derived From the Balloons Game, and Correlations Among Them

	Mean (*SD*)	ER negative	ER positive	ER failure	ER success	ER total
ER negative	–.62 (.27)	—				
ER positive	–.73 (.47)	$.16^{+}$	—			
ER failure	–.66 (.23)	$.55^{***}$	$.49^{***}$	—		
ER success	–.65 (.24)	$.51^{***}$	$.58^{***}$	$.47^{***}$	—	
ER total	–.66 (.20)	$.63^{***}$	$.57^{***}$	$.84^{***}$	$.83^{***}$	—

$^{+}$ $p < .10$. *** $p < .001$.

Discussion

The aim of this research was to formally derive and experimentally extract indices that are proportional to regulatory process in emotion-expressive behavior. The results indicate that the measures of emotional-expressive behaviors and emotion expression regulation derived from the Balloons Game have internal reliability and conceptual validity. The primary manipulation of the Balloons Game, the instruction to restrain facial expressions of emotion, led to a significant reduction in emotion-expressive behaviors, indicating that 6-year-olds understand and are able to adjust their performance in compliance with task demands. Significant intertrial correlations indicate reasonable stability of the measures, supporting the general internal reliability of the instrument. Test-retest stability of the emotion regulation indices over a more extended time period was not tested in the present study. Collection of such data could be a valuable contribution of future studies using the Balloons Game.

Overall, the correlations between indices of regulation of emotion-expressive behavior across all conditions suggest the involvement of a common core process controlling conscious and intentional down-regulation of emotion. However, this interpretation should be taken with caution because at least one correlation (between regulation of negative and positive facial expressions) was low and appeared at a trend level of significance. This suggests that these two regulatory processes share little variance. This finding may be in accord with research indicating relative independence, rather than continuity of positive and negative affect (Diener & Emmons, 1984; Tellegen, Watson, & Clark, 1999; Watson, Clark, & Tellegen, 1988). Additional studies are needed to verify the nature of associations between the regulation of positive and negative emotion expressions and possibly identify additional factors influencing these associations. Further, identifying the neural architecture supporting the regulatory processes indexed by the present task is of great interest and may be achieved through adaptation of the Balloons Game to imaging research contexts. Our data provide preliminary hints for limbic involvement through the associations found here between heart period and the emotion regulation indices. Future studies may wish to use more sophisticated variables, such as heart rate variability (Berntson et al., 1997; Porges, Doussard-Roosevelt, Portales, & Greenspan, 1996), and record online during task performance, rather than at rest.

An important aim of the present study was to provide external validation to the Balloons Game as a measure of emotion regulation. We focused on measurement of effortful suppression of emotion-expressive behavior as one form of emotion regulation. Therefore, for the purpose of external validation, we selected measures that could potentially represent other aspects of emotion regulation. The significant, meaningful correlations found between the emotion regulation indices derived from the Balloons Game and the external indices of externalizing behavior problems, effortful control, attentional focusing, activity level, emotional tone of mother-child interaction, and heart rate provide strong evidence that our new task, indeed, taps into the larger construct of emotion regulation. Further, this validation is based on multiple sources of information (e.g., parental reports, behavioral observations, and psychophysiology). As such, our conceptual framework and the Balloons Game may be viewed as an additional step in the theoretical and empirical effort to elucidate measurement that is truly proportional to a regulatory process.

The derivation of behavioral raw expressivity scores alongside regulatory indices allows for a systematic assessment of these processes in the context of developmental and clinical research. Future studies could use the Balloons Game and its conceptualization to explore mediating and moderating factors, such as socialization, culture, cognitive development, and peer relationships on the maturation of emotion regulatory capacities. Clinical and etiological research could now estimate the distinct contributions of emotional expressivity and emotion regulation to psychiatric conditions, such as conduct disorder, affective disorders, and anxiety.

Here, we focused on an experimental paradigm that allows the derivation of the estimate $(1 - R)/(1 - nR)$, that is proportional to emotion regulation (see Equations 1, 2, and 3 above). Potentially, a similar procedure may also be used to derive an estimate that is proportional to raw emotional expressivity. Thus, regulation instructions may be kept constant, whereas the intensity of emotional provocation is manipulated experimentally. For instance, success may be thwarted while a child works toward a highly attractive versus a less-attractive prize. It is important to note, however, that

Table 3
Correlations Between Raw Expressivity Scores During Low- and High-Regulation Demands and the Indices of Emotion Regulation (ER) Derived From the Balloons Game

	ER negative	ER positive	ER failure	ER success	ER total
Negative expressivity					
Low regulation	–.16	–.13	–.09	–.15	–.16
High regulation	$-.57^{***}$	$-.33^{***}$	$-.39^{***}$	$-.40^{***}$	$-.48^{***}$
Positive expressivity					
Low regulation	.03	$-.18^{*}$	–.13	–.12	–.18
High regulation	$-.21^{*}$	$-.55^{***}$	$-.43^{***}$	$-.43^{***}$	$-.53^{***}$
Expressivity during failure					
Low regulation	–.14	–.15	$-.19^{*}$	–.07	$-.18^{*}$
High regulation	$-.45^{***}$	$-.47^{***}$	$-.61^{***}$	$-.31^{***}$	$-.55^{***}$
Expressivity success					
Low regulation	–.04	–.17	–.06	$-.21^{*}$	–.15
High regulation	$-.29^{**}$	$-.52^{***}$	$-.30^{**}$	$-.67^{***}$	$-.56^{***}$
Total expressivity					
Low regulation	–.07	$-.20^{*}$	–.15	–.15	–.20
High regulation	$-.45^{***}$	$-.59^{***}$	$-.55^{***}$	$-.55^{***}$	$-.66^{***}$

* $p < .05$. ** $p < .01$. *** $p < .0001$.

Table 4
Means and Standard Deviations of the External Indices of Emotion Expression and Regulation and Correlations Between the Emotion Regulation Indices Derived From the Balloons Game and These External Indices

	M (*SD*)	ER negative	ER positive	ER failure	ER success	ER total
CBCL: externalizing	7.88 (6.45)	$-.32^{***}$	$-.21^{*}$	$-.20^{*}$	$-.22^{*}$	$-.27^{**}$
CBCL: internalizing	6.16 (6.00)	.00	–.10	.00	–.05	–.05
CBQ: activity level	3.99 (.99)	$-.17^{+}$	$-.27^{**}$	$-.16^{+}$	$-.26^{**}$	$-.27^{**}$
CBQ: inhibitory control	5.39 (.93)	.11	$.29^{***}$	.14	$.23^{**}$	$.22^{**}$
CBQ: attentional focusing	5.31 (.83)	.09	$.20^{*}$	.12	.15	$.18^{*}$
Quality of M-C interaction	3.74 (.73)	$.24^{**}$	.14	$.16^{+}$	$.15^{+}$	$.21^{*}$
Heart Period	671.44 (92.18)	–.09	$-.22^{*}$	$-.24^{**}$	$-.22^{*}$	$-.22^{*}$

$^{+}p < 0.10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

a formal description of emotional expressivity is beyond the scope of the present investigation. It is also important to point out that although the ratio index of emotion regulation used in the present study has many advantages, in certain circumstances, the simple difference between high and low regulatory efforts can be instructive. In fact, there are various ways in which the rich data set derived from the Balloons Game could be analyzed, and researchers are encouraged to use the metric that best reflects the variables of interest in their specific study. For instance, a revised equation could weigh the ratio index by the difference between the high- and low-regulation conditions.

Some limitations of the present theory and study should be further acknowledged and considered. First, we studied regulation of emotion expressive behavior, which is not the same as regulation of subjective emotional experience. It is important to acknowledge that there are instances in life in which the expressed behavior and the subjective experience of emotion are incongruent. Although such instances are more common in adults than in young children, this possibility should be considered when interpreting our finding of regulated expressed behavior. Second, and related to the preceding point, if expressive behavior and subjective emotional experience are, in fact, related, then our instruction to regulate (i.e., suppress) facial expressions might have changed not only the regulatory component of the observed behavior, but also the level of experienced emotion (Gross & Levenson, 1997). This possibility could be assessed in future studies by asking the participants to report on their subjective experience in the Balloons Game. Finally, we measured children's ability to suppress facial expressions of emotion within a laboratory procedure explicitly asking them to do so. Although correlations between our ER indices and parental reports on their children's typical styles of emotion regulation did emerge, our data do not address associations between our laboratory-obtained findings and behavior in everyday life. Children might have been compliant and masked their emotion in the lab, but this does not necessarily mean that they are similarly compliant or capable in other contexts. This is an important issue that could be addressed in future studies.

To conclude, deriving an index of emotion regulation that could be applied in a within-subject context is a challenging task, both theoretically and empirically. It is, however, essential to pursue this goal if, eventually, indices of emotion regulation are to be applied and measured as factors of individual differences. Clearly, the procedures targeting this task are based on a set of certain assumptions. The fact that researchers do not always make explicit these assumptions does not take away from the fact that their experimental procedures, data analyses, and interpretation are, nevertheless, dependent on them. We concur with Cole et al. (2004) that with the present state of affairs in emotion regulation research, a good strategy would be to formally present the assumptions underlying study procedures so that these could be subjected to scientific scrutiny. It is our hope that applying the conceptual framework we laid out here to the experimental study of emotion regulation would make a valuable contribution for testing hypotheses and for discussion about emotion regulation processes as an individual differences factor.

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist 4*–18. Burlington, VT: University of Vermont, Department of Psychiatry.

Allhusen, V., Belsky, J., Booth, C. L., Bradley, R., Brownell, C. A., Burchinal, M., . . . , Weinraub, M. (2004). Affect dysregulation in the mother-child relationship in the toddler years: Antecedents and consequences. *Development and Psychopathology, 16,* 43–68.

Apter, A., Offer, R., Blumenson, R., & Tyano, S. (1988). Use of the Child-Behavior Checklist in an Israeli Adolescent Psychiatric Unit. *Acta Psychiatrica Scandinavica, 78,* 96–101.

Barlow, D. H. (1991). Disorders of emotion. *Psychological Inquiry, 2,* 58–71.

Berntson, G. G., Bigger, J. T., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., . . . , van der Molen, M. W. (1997). Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology, 34,* 623–648.

Byrne, E. A., & Porges, S. W. (1993). Data-dependent filter characteristics of peak valley respiratory sinus arrhythmia estimation—a cautionary note. *Psychophysiology, 30,* 397–404.

Calkins, S. D., & Fox, N. A. (2002). Self-regulatory processes in early personality development: A multilevel approach to the study of childhood social withdrawal and aggression. *Development and Psychopathology, 14,* 477–498.

Calkins, S. D., Graziano, P. A., Berdan, L. E., Keane, S. P., & Degnan, K. A. (2008). Predicting cardiac vagal regulation in early childhood from maternal-child relationship quality during toddlerhood. *Developmental Psychobiology, 50,* 751–766.

Calkins, S. D., & Johnson, M. C. (1998). Toddler regulation of distress to frustrating events: Temperamental and maternal correlates. *Infant Behavior & Development, 21,* 379–395.

Cole, P. M. (1986). Children's spontaneous control of facial expression. *Child Development, 57,* 1309–1321.

Cole, P. M., Martin, S. E., & Dennis, T. A. (2004). Emotion regulation as a scientific construct: Methodological challenges and directions for child development research. *Child Development, 75,* 317–333.

Cole, P. M., Teti, L. O., & Zahn-Waxler, C. (2003). Mutual emotion regulation and the stability of conduct problems between preschool and early school age. *Development and Psychopathology, 15,* 1–18.

Cole, P. M., ZahnWaxler, C., Fox, N. A., Usher, B. A., & Welsh, J. D. (1996). Individual differences in emotion regulation and behavior problems in preschool children. *Journal of Abnormal Psychology, 105,* 518–529.

Davidson, R. J., Jackson, D. C., & Kalin, N. H. (2000). Emotion, plasticity, context, and regulation: Perspectives from affective neuroscience. *Psychological Bulletin, 126,* 890–909.

Derryberry, D., & Rothbart, M. K. (1997). Reactive and effortful processes in the organization of temperament. *Development and Psychopathology, 9,* 633–652.

Diener, E., & Emmons, R. A. (1984). The independence of positive and negative affect. *Journal of Personality and Social Psychology, 47,* 1105–1117.

Eisenberg, N., Cumberland, A., Spinrad, T. L., Fabes, R. A., Shepard, S. A., Reiser, M., . . . , Guthrie, I. K. (2001). The relations of regulation and emotionality to children's externalizing and internalizing problem behavior. *Child Development, 72,* 1112–1134.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial action coding system (FACS)—the manual.* Salt Lake City, UT: A Human Face.

Falender, C. A., & Heber, R. (1975). Mother-child interaction and participation in a longitudinal intervention program. *Developmental Psychology, 11,* 830–836.

Fox, N. A. (1994). Dynamic cerebral processes underlying emotion regulation. In N. A. Fox (Ed.), *The development of emotion regulation: Biological and behavioral considerations*. Serial no. 240. *Monographs of the Society for Child Development, 59,* 2–3.

Frijda, N. H. (1988). The laws of emotion. *American Psychologist, 43,* 349–358.

Giuliani, N. R., McRae, K., & Gross, J. J. (2008). The up- and down-regulation of amusement: experiential, behavioral, and autonomic consequences. *Emotion, 8,* 714–719.

Gross, J. J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology, 74,* 224–237.

Gross, J. J. (1999). Emotion regulation: Past, present, future. *Cognition & Emotion, 13,* 551–573.

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology, 39,* 281–291.

Gross, J. J., & Levenson, R. W. (1993). Emotional suppression: Physiology, self-report, and expressive behavior. *Journal of Personality and Social Psychology, 64,* 970–986.

Gross, J. J., & Levenson, R. W. (1997). Hiding feelings: The acute effects of inhibiting negative and positive emotion. *Journal of Abnormal Psychology, 106,* 95–103.

Gross, J. J., & Munoz, R. F. (1995). Emotion regulation and mental-health. *Clinical Psychology Science and Practice, 2,* 151–164.

Kagan, J. (1994). On the nature of emotion. In N. A. Fox (Ed.), *Monographs of the Society for Research in Child Development: Vol. 59 (2–3).* The development of emotion regulation: Biological and behavioral considerations (pp. 7–24). Chicago, IL: University of Chicago Press.

Kopp, C. B. (1989). Regulation of distress and negative emotions: A developmental view. *Developmental Psychology, 25,* 343–354.

Oppenheim, D., Nir, A., Warren, S., & Emde, R. N. (1997). Emotion regulation in mother-child narrative co-construction: Associations with children's narratives and adaptation. *Developmental Psychology, 33,* 284–294.

Porges, S. W., Doussard-Roosevelt, J. A., Portales, A. L., & Greenspan, S. I. (1996). Infant regulation of the vagal "brake" predicts child behavior problems: A psychobiological model of social behavior. *Developmental Psychobiology, 29,* 697–712.

Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology, 12,* 427–441.

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.

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The children's behavior questionnaire. *Child Development, 72,* 1394–1408.

Rubin, K. H., Coplan, R. J., Fox, N. A., & Calkins, S. D. (1995). Emotionality, emotion regulation, and preschoolers social adaptation. *Development and Psychopathology, 7,* 49–62.

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

Shipman, K. L., & Zeman, J. (2001). Socialization of children's emotion regulation in mother-child dyads: A developmental psychopathology perspective. *Development and Psychopathology, 13,* 317–336.

Stansbury, K., & Gunnar, M. R. (1994). Adrenocortical activity and emotion regulation. In *The development of emotion regulation: Biological and behavioral considerations. Monographs of the Society for Research in Child Development, 59,* 108–134.

Tellegen, A., Watson, D., & Clark, L. A. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science, 10,* 297–303.

Thompson, R. A. (1991). Emotional regulation and emotional development. *Educational Psychology Review, 3,* 269–307.

Thompson, R. A. (1994). Emotion regulation: A theme in search of definition. In *The Development of emotion regulation: Biological and behavioral considerations*. *Monographs of the Society for Research in Child Development, 59,* 25–52.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology, 54,* 1063–1070.

Zilber, N., Auerbach, J., & Lerner, Y. (1994). Israeli norms for the Achenbach Child-Behavior Checklist: Comparison of clinically-referred and non-referred children. *Israel Journal of Psychiatry and Related Sciences, 31,* 5–12.

Received February 9, 2009
Revision received March 22, 2010
Accepted April 19, 2010 ■