

The development of adaptive risk taking and the role of executive functions in a large sample of school-age boys and girls

Morris D. Bell^{a,b,*}, Ahmet Esat Imal^a, Brian Pittman^a, Grace Jin^a, Bruce E. Wexler^a

^a Yale University School of Medicine, 300 George Street, New Haven, CT 06511 United States

^b VA Connecticut Healthcare System, 950 Campbell Avenue, West Haven, CT 06516 United States

ARTICLE INFO

Keywords:

Self-regulation
Balloon analogue risk task (BART)
Bubblegum analogue risk task for children (BART-C)
Adaptive risk-taking
Executive function
Cognitive development

ABSTRACT

Background: The Balloon Analogue Risk Task for Children (BART-C) demands self-regulation of emotion that requires risk-tolerance and adaptive risk-taking to make good decisions under stress (hot cognition).

Methods: BART-C measures of adaptive risk-taking in 5,409 children K-8th grade were analyzed for improvements by grade, for relationships to executive functioning (EF) and for associations with school characteristics and academic achievement.

Findings: BART-C improved across grades. Boys showed significantly greater Recklessness, particularly in middle school. EF was a partial mediator between grade and Variability and Recklessness. Better BART-C Total score and less Recklessness were related to lower free-or-reduced-school-lunch percentage and better math and reading proficiency of children's schools.

Conclusions: BART-C is a potential "hot-cognition" measure of self-regulation and adaptive risk-taking for children.

1. Introduction

Self-regulation of emotions and behavior has been long recognized as a core feature of adaptive development from infancy [1] into adult life. It requires the ability to monitor and control one's behavior, emotions, or thoughts, altering them according to the situation. To do so involves regulating emotions, inhibiting first responses, sustaining attention despite irrelevant stimuli and making adaptive behavioral decisions to reach one's goal [2–4]. Self-regulation in children has been shown to predict academic achievement and resilience to adverse circumstances; and a growing body of literature indicates that students who can self-regulate cognitive, motivational, and behavioral functioning are more effective as lifelong learners [5–7]. Children's self-regulatory difficulties have been associated with poverty-related stress, perhaps because they live in environments that do not effectively promote development of attention skills, following rules, or controlling impulses [7]. A child's capacity for self-regulation is related to their ability to cope with adverse experiences during childhood such as serious illness, loss, trauma, divorce and family dysfunction [8,9]. Difficulty with emotional self-regulation and tolerating distress has also been linked to depression, antisocial behavior [10], addiction [11], self-harm and suicidal ideation [12] in later life.

Because of its importance, self-regulation has been an aim of early

intervention programs such as Tools of the Mind [13] and the Rochester Resilience Project [14], and findings have supported the relative plasticity of self-regulation through enriched experiences that emphasize executive function training [4]. Executive functions (EF) are linked to neurocircuits involving the prefrontal cortex, which develop throughout childhood and adolescence and include abilities such as sustained attention, inhibitory control, working memory, and cognitive flexibility [7]. Executive functions have a top-down influence on emotions, but emotional states that are accompanied by physiological changes can overwhelm the top-down influences of the prefrontal cortex [15–17]. The bottom-up influence of these emotional states is particularly common for younger children whose executive function is just developing.

Scientific exploration of self-regulation depends on having research tools that capture this construct. Rating scales completed by parents and teachers (e.g. Child Behavior Rating Scale [18]; The Behavior Rating Inventory of Executive Functions [19]; Child Behavioral Checklist [20]) provide systematic reports on behaviors related to executive functioning and self-regulation, but these are functional reports and not performance tests that provide a direct measure of self-regulation in action. Performance measures commonly used in self-regulation studies rely on cognitively demanding tasks that are not usually emotionally evocative. These include cognitive tests of EF that measure

* Corresponding author.

E-mail address: morris.bell@yale.edu (M.D. Bell).

<https://doi.org/10.1016/j.tine.2019.100120>

Received 22 March 2019; Received in revised form 9 September 2019; Accepted 9 October 2019

Available online 14 October 2019

2211-9493/ © 2019 Elsevier GmbH. All rights reserved.

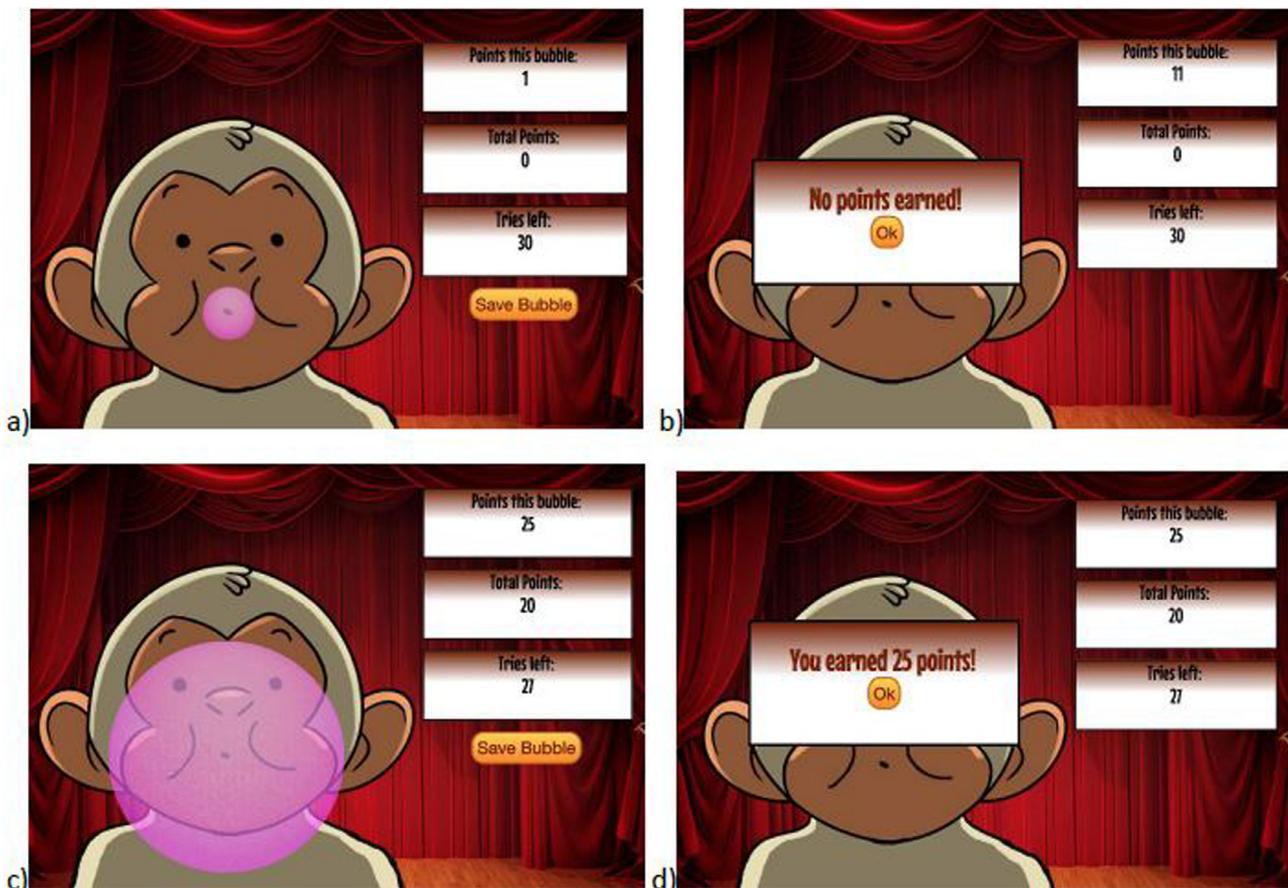


Fig. 1. Depicted here is the BART-C “Great Bubble Gum Contest”: a) First trial after 1 tap. b) First trial, an explosion occurred after 11 taps and a “No points earned” message was displayed. c) Fourth trial, with 20 total earned points and 25 current points accumulated. d) Fourth trial, participant earned 25 points after clicking “Save Bubble”.

response inhibition, working memory and attention (e.g. NIH Toolbox of Executive Function) and a few tasks involving movement such as the Head-Toes-Knees-Shoulders Task (HTKS), developed for children ages 4–6, which directly assesses self-regulation using an opposites task (“touch your head when you hear toes”; [21]).

An often-overlooked aspect of this type of performance testing is that these tasks are not especially designed to evoke emotions. A commonly used term in this regard is “cold cognition”, which implies that the task does not include emotional arousal and that good performance does not require managing emotions. Although the child may become frustrated or discouraged while doing such a task, the task itself is not constructed to arouse negative feelings. In contrast, “hot cognition” is cognitive performance that requires emotion regulation. Since self-regulation in daily life usually combines reasoning and emotional control, an executive function task that involves “hot cognition”, might come closer to the underlying construct of self-regulation than do current “cold cognition” measures such as a Go/No-Go response inhibition task or a Stroop Task [22]. With these considerations in mind, we created a child adaptation of the Balloon Analogue Risk Task for Youth (BART-Y; [23]) which is a well-validated, computer-based task of risk-taking and which was recently linked to executive function in adolescents [24].

1.1. Balloon analogue risk task (BART) as a measure of adaptive risk taking

We chose the BART because it was developed specifically to be an emotionally evocative risk-taking task (like gambling) used in studies of sensation-seeking, impulsivity and delayed discounting. On a computer

screen, participants are presented with a small balloon next to buttons labeled Pump and Collect \$\$\$\$. Each click or Pump incrementally inflates the balloon, adding 5 cents to a temporary reserve. At any point, the participant can Collect \$\$\$ and transfer money to a “Total Earned” reserve bank. Each balloon can explode after any number of pumps, resulting in no money earned. A new uninflated balloon appears after each explosion or point collection. The BART variable for measuring risk-taking propensity is Adjusted Pumps, the number of inflations that did not explode. This index is preferred because it selects trials where risk-taking behavior is unconstrained by the explosion point [25]. In adults, BART Adjusted Pumps has been related to self-reports of risk-taking propensity with high scores linked to alcohol use, delinquency, unprotected sex, smoking [26], marijuana use [27], crack cocaine use [28], conduct disorders [29], and risky behavior in inner-city adolescents [30].

While the BART has mostly been used to measure risk-propensity as maladaptive behavior, it may also be used to measure adaptive risk-taking. This alternative interpretation of BART scores as possibly being a measure of self-regulation of emotion in the interest of goal achievement first came to our attention when one of our authors (MDB) analyzed pre-post BART data from a study he conducted of a cognitive training intervention for adults with substance use disorders (SUD) [31]. He had included the BART with the hypothesis that better substance abuse outcomes would be related to reduced risk-taking. Contrary to this hypothesis, increased risk-taking (Adjusted Pumps) and increased Total Score were associated with better SUD outcomes and related to higher scores on executive functioning at baseline. Although never published, this unexpected finding was in accord with an alternative explanation of BART data reported by DeMartini et al. [32] who

showed that higher Coefficient of Variability (COV), calculated as the standard deviation of adjusted inflations divided by the mean of adjusted puffs, among young adult heavy drinkers indicated adaptive risk-taking. COV was associated with lower lifetime drinking quantity and higher levels of self-efficacy to control drinking.

Most recently, Blair et al. [24] published a study of the BART with 105 adolescents (mean age = 17.91) utilizing the same COV measure of adaptive risk-taking. They found that adolescents with less intra-individual variability (COV) earned better scores, that COV improved with age, that working memory predicted COV, and that working memory mediated the relationship between age and COV. They suggested that working memory “may allow adolescents and young adults to more effectively process information to navigate uncertain options during risk-taking that ultimately leads to advantageous outcomes.”

1.2. BART-C: A measure of adaptive risk-taking for children

The present study continues the exploration of the BART as a “hot cognition” measure of self-regulation and adaptive risk-taking with a much younger and larger sample. For this purpose, we adapted the BART to make it more suitable for children, using an image of a monkey blowing up bubblegum (Fig. 1). The original BART was for ages 18 years and above and a Youth version was created by the originators of the BART because they thought that it needed to be adapted for adolescents. Similarly, we adapted BART for children by making it a little shorter and with the more playful appearance of the monkey. It uses the same point system, but we altered the random explosion rate to include an explosion after two “puffs” within the first five trials. We did so to insure early exposure to an aversive outcome from which the child must recover to perform well. We used Total Score, Adjusted Puffs and the COV measure for each child, and added an additional measure of Recklessness. Because we had a sufficiently large sample to create grade-adjusted norms for our measures, we could relate the degree of risk relative to their peers (Adjusted Puffs) that each child was willing to take to the Total score that they obtained, relative to their peers. We did this by subtracting grade-adjusted Z-score for Total Score from grade adjusted Z-score for Adjusted Puffs. Children who took greater risks but obtained lower Total score compared to their peers, we regard as “reckless”. We named our test the Bubblegum Analogue Risk Task for Children (BART-C) to acknowledge it as a version of BART, while not wanting to ignore its minor differences.

Using these measures, we hypothesized that 1) BART-C performance on all measures would improve cross-sectionally by grade as would be expected with normal development of self-regulation. 2) BART-C performance measures would show moderate relationships with measures of EF, since these abilities underlie self-regulation as described in the literature cited above. 3) The relationship between BART-C measures and grade would be mediated by EF measures as was reported in a previous study with adolescents [24]. And, 4) Better BART-C performance would be inversely related to the schools’ poverty level (percent free or reduced school lunch) and positively related to academic ratings, since self-regulation has been previously associated with these variables as described in the literature cited above [7,8,9].

2. Materials and methods

2.1. Participants

Our study utilizes archival BART-C data from children participating in ACTIVATE, a computerized cognitive training program designed by C8 Sciences (<https://www.c8sciences.com>). The C8 Sciences ACTIVATE program was developed by Bruce Wexler, MD, Professor of Psychiatry at Yale University, as a method for enhancing neurocognitive functioning in children and is commercially available for school and home use. At baseline, ACTIVATE delivers a suite of online EF tests to children in their classroom. In our sample, 5409 children between grades K-

8 were administered the BART-C. Our large sample represents a diverse population from 199 educational institutions, 40 US states, and 16 countries. There is no personal identifying information in these records; data are archived for product improvement, and this study is exempt from IRB review under Category 2 for educational and cognitive tests with no personally identifying information.

2.2. BART-C

BART-C was designed by C8 Sciences. In Fig. 1, our “Great Bubble Gum Contest” begins by displaying a monkey on a computer screen. With each tap on the space bar or mouse, the monkey blows up the bubble, with more puffs generating more points. Each bubble can go up incrementally to 50 puffs but can explode at any point. During each trial, the child can stop enlarging the bubble and click “Save” to add points to the Total Score, but if the bubble bursts, all the earned points are lost.

There are 30 trials. The optimal number of puffs per trial is 38, but few children will consistently go that high. While the explosion point is random, an explosion at 2 puffs always occurs within the first 5 trials. This startles players and usually causes them to be cautious. A few never recover and won't go beyond 2 puffs, but most will cautiously increase their number of puffs across trials. Across trials, players who are not overwhelmed by their emotions will learn that risk-tolerance and consistent play (ie. less variability) leads to a better score.

2.3. BART-C variables

BART-C variables include Total Score, which is the cumulative number of saved points; Adjusted Puffs, which is the average number of bubblegum puffs when the bubble didn't explode; and two additional measures – the Coefficient of Variability (COV [32]), which is the standard deviation of Adjusted Puffs divided by the mean of Adjusted Puffs, and our Recklessness measure, which is derived by subtracting the grade-adjusted Z-score for Total Score from the grade-adjusted Z-score for Adjusted Puffs. Thus, higher Recklessness occurs when a child risks more than his peers while earning less than his peers, while low Recklessness occurs when a child risks less than her peers but earns more than her peers. For example, a child might receive a Total Score of 200 by consistently puffing 50 times on each trial (Adjusted Puffs = 50) and having only 4 successful trials ($4 \times 50 = 200$). That would produce a very high Z-score for Adjusted Puffs while producing an average Z-score for Total Score. Subtracting the Z-score for Total Score from the Z-score for Adjusted Puffs will produce a highly positive Recklessness score, indicating much greater risk-taking compared to reward. Another child might consistently use 20 puffs (Adjusted Puffs = 20) and have 10 successful trials ($20 \times 10 = 200$) receiving the same Total Score as the child who used 50 puffs. This child's Recklessness score would likely be close to 0 because the Z-score for Adjusted Puffs and Z-score for Total Score would be about the same. A third child is inconsistent in the number of puffs across trials and ends up with an Adjusted Puffs score of 20 (average for his grade) but a Total Score of only 100 points (below average for his grade). Even though the Adjusted Puffs Z-score is the same as the child in the previous example, the Total Score Z-score is much lower. This child will then have a positive Recklessness score because the child took greater risks than the reward received.

To further illustrate the third example, imagine a driver who generally drives at the same rate as other drivers (Adjusted Puffs), except that in moments of anger, excitement or frustration, the driver floors it! The driver gets pulled over by the police or gets into an accident whenever that happens (burst bubbles). Over 30 trips (Trials), the driver's average miles travelled without getting stopped (Total Score) is far below average compared with his usual average rate of speed (Adjusted Puffs). This driver's insurance company would likely conclude that he is a “reckless driver.” The term Recklessness means to act without regard to consequences. It is maladaptive risk-taking; and this

simple formula of subtracting the Z score for Total Score from the Z score for Adjusted Puffs is our effort to measure that construct.

2.4. Executive function (EF) measures

The Flanker Focused Attention Task and the List Sorting Working Memory Test closely follow the test specifications in the NIH Toolbox of Executive Function. For the Flanker Task, students see five arrows in a row, the two arrows to the right of the middle arrow and the two arrows to the left of the middle arrow are pointing in the same direction. Using the arrow keys on their keyboard, the students need to identify which way the arrow in the middle is pointing. The middle arrow may be pointing in the same direction (congruent trial) or in the opposite direction (incongruent trial) of the arrows on either side. To perform the task accurately and quickly, students must mentally ignore the flanking arrows. Performance scores include accuracy and reaction time for congruent and incongruent trials. The primary measure is average Reaction Time on Incongruent Trials.

The List Sorting Working Memory Test (WMT) presents the child with a series of animals or household objects. The child then selects the objects just seen from among a grid of 16 objects, clicking them in order from smallest to largest rather than the order in which they were presented. The test starts with a list of 2 objects, and list length is increased by one for accurate responding. If the child fails twice at the same list length, the task ends. In part one, trials of animals and household objects alternate. In part two, animals and household objects are presented in the same trial, and the child must reorder the animals first and then the household objects. As a validity check, if a child was unable to report back two items in correct order, we considered it possible that the child did not understand or engage with the test, and the data were excluded from analysis.

The primary measure was sum of correct list lengths. Flanker and WM are available at: <http://www.healthmeasures.net/explore-measurement-systems/nih-toolbox/intro-to-nih-toolbox/cognition>

The Go/No-Go test of Response Inhibition (GNG) instructs the child to press the space bar whenever a “Go” stimulus is presented but not when a “No-Go” stimulus is presented. There are three blocks with different stimuli, 50 stimuli per block with 40 Go and 10 No-Go trials, randomized in sets of 10 with 8 Go and 2 No-Go in each set. In the first block “P” is the Go stimulus and “R” is the No-Go stimulus. In the second block, this is reversed. In the third block, pictures of furniture are the Go trials and pictures of foods like cake and ice cream are the No-Go stimuli. Tests with less than 85% correct response to Go-Trials were regarded as invalid because the child failed to establish the consistent response bias required to measure response-inhibition. Only children who were in school for all test days and had valid data on all three tests were included in the study sample.

2.5. Data cleaning and analysis

Children were included in the BART-C data set if their performance met embedded validity criteria. For a child's BART-C to be considered valid it had to have a Total Score and Adjusted Score > 1, and no more than one trial where the child hit the save button before any puffs occurred. The cleaned dataset was used to create whole-sample Z-scores for each variable. These z-scores were then converted to standard scores (mean=100, SD=15) allowing for a direct examination of grade-related mean changes for all BART-C measures. Whole-sample Z-scores were also generated for EF measures. Only participants with valid scores on all measures ($N = 4200$) were used for hypotheses related to EF. BART-C measures among participants from schools in the USA were aggregated at the school level for correlations with school characteristics, which were obtained from the US Department of Education statistics aggregated by Niche (www.niche.com). We included only those grades with 20 children or more.

To account for the hierarchical/correlated nature of the data – i.e.,

grade within school – mixed-effects regression [33] using PROC MIXED in SAS was used to characterize trajectories of BART-C measures across grade levels using equivalent standard scores as the outcome. Each model included grade as a fixed effect and school as a random effect. In separate models, grade was first considered as a categorical predictor and then as a continuous predictor to test for polynomial trends over grade. Main and interactive effects of gender were also considered in each model followed by appropriate post-hoc contrasts. Correlation analysis was used to test for associations between BART-C variables and measures of EF and between BART-C Z-score means for each grade and the grade's Niche school characteristics.

We tested EF as a mediator of the relationship between grade and BART-C COV and Recklessness Z-scores following the mediation procedure as outlined by Baron and Kenny [34]. Within this cohort ($n = 4200$), mediation analysis was conducted across all grades (K-8) for COV as a linear model fit the data well ($b_{\text{grade k-8}} = -0.07$, $p < .0001$). For Recklessness, a quadratic model fit the data well ($b_{\text{grade}^* \text{grade}} = 0.013$, $p < .0001$), driven by a linear decline from kindergarten through 4th grade ($b_{\text{grade k-4}} = -0.087$, $p < .0001$) and then little change thereafter ($b_{\text{grade} > 4} = 0.022$, $p < .27$). We therefore restricted the mediation analysis for Recklessness to grade K-4. Each regression model accounted for the hierarchical nature of the data by modeling random school effects as above. The Sobel test [35] was used to test for significance of indirect effects. In lieu of testing the mediating effects of each EF measure separately, for statistical parsimony factor scores were estimated based on a one-factor solution for the Flanker, GNG, and WMT tasks. Specifically, a principle component factor analysis using PROC FACTOR in SAS yielded a one factor solution from which overall EF scores were calculated by multiplying individual scores by optimally determined weights for each variable and then summing the products. All analyses assumed normal distribution, and alpha was set at 0.05 and were conducting using SAS, version 9.4 (Cary, NC).

3. Results

3.1. Cross-sectional developmental analysis of BART-C measures

Estimated least-square means and standard errors estimated from the mixed model for each BART outcome are depicted in Fig. 2. BART total standard scores increased gradually in a linear fashion ($b_{\text{grade k-8}} = 1.3$, $p < .0001$) across grades K-8. A quadratic model fit BART adjusted scores well ($b_{\text{grade}^* \text{grade}} = 0.14$, $p = .002$) with little observed increase in scores up to grade 3 ($b_{\text{grade k-3}} = -0.10$, $p = .76$) followed by a linear increase after 3rd grade ($b_{\text{grade} > 3} = 1.3$, $p < .0001$). A quadratic model fit the Recklessness scores well ($b_{\text{grade}^* \text{grade}} = 0.22$, $p < .0001$), driven by a linear decline from kindergarten through 4th grade ($b_{\text{grade k-4}} = -1.4$, $p < .0001$) and then little change thereafter ($b_{\text{grade} > 4} = 0.32$, $p < .26$). A linear model fit the COV well ($b_{\text{grade k-8}} = 1.3$, $p < .0001$) with a steady decrease from kindergarten through 8th grade. These findings affirm our first hypothesis that adaptive risk-taking performance (Total Score), risk-tolerance (Adjusted Puffs) and self-regulation (COV and Recklessness) improves concurrently with age, although with different patterns for different measures.

Gender by grade interactions were observed for both BART adjusted puffs ($p = .0005$) and Recklessness ($p = .005$). As shown in Figs. 3a and 3b, males began taking greater risk (adjusted puffs) and exhibiting greater recklessness compared to their female counterparts around 5th grade with steady divergence thereafter. Post hoc gender contrasts for each BART measure are shown in Table 1. Despite some gender differences in earlier grades, overall patterns across grades were similar between males and females for both BART total ($p = .13$) and COV ($p = .33$).

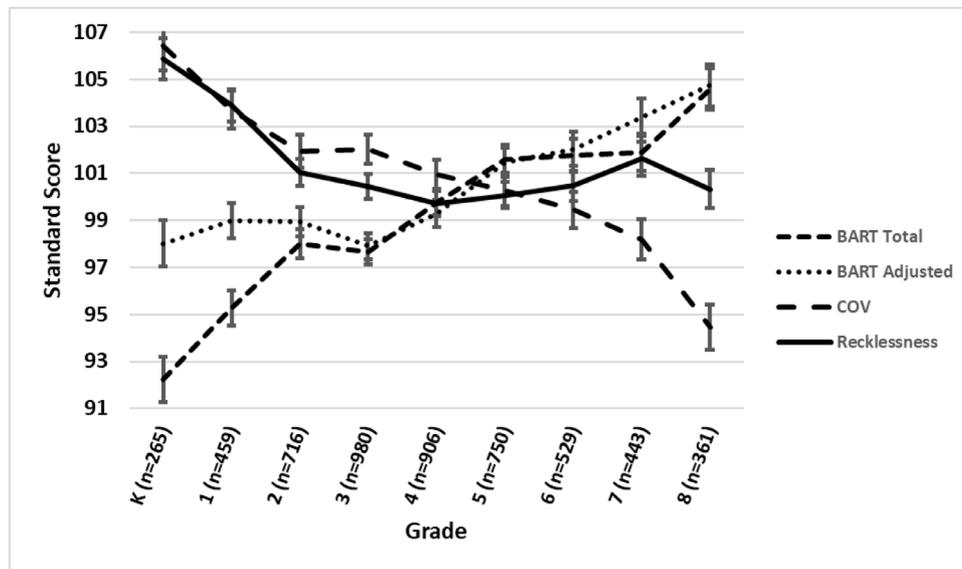


Fig. 2. Standard Scores for each BART measure across grade.
 Note: Values are least-square means (LS means) and standard errors estimated from mixed models.

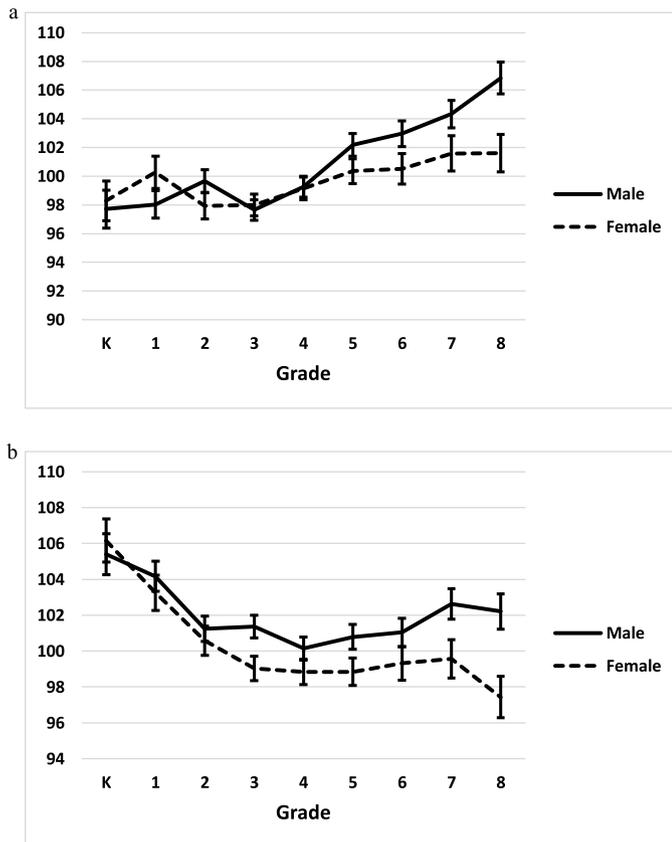


Fig. 3. a). BART Adjusted Puffs Standard Scores by Gender across Grade.
 Note: Values are least-square means (LS means) and standard errors estimated from the mixed model.
 b). BART Recklessness Standard Scores by Gender across Grade.
 Note: Values are least-square means (LS means) and standard errors estimated from the mixed model.

3.2. Mediation model of EF

A series of regression models were used to test the hypothesis that EF mediates the observed effects of grade on both COV ($b = -0.071$,

$p < .0001$) and Recklessness ($b = -0.087$, $p < .0001$). Analyses examining the mediating effects of EF on the association between grade and Recklessness was restricted to grades K through 4th grade, after which there is no longer a linear relationship (See Fig. 2). Results of the mediation analyses are illustrated in Fig. 4. The association between grade and EF was significant across all grades ($b = 0.239$, $p < .0001$) and grades K – 4th ($b = 0.329$, $p < .0001$), as was the effect of EF on COV ($b = -0.081$, $p < .0001$) and Recklessness ($b = -0.059$, $p = .004$) after controlling for grade, supporting the mediation hypothesis. The estimated indirect effect was $(0.239) \times (-0.081) = -0.019$ on COV and $(0.329) \times (-0.059) = -0.019$ on Recklessness, with both significant ($p < .0001$) according to the Sobel test [34]. The effect of grade, after controlling for EF, was significant, but diminished on both COV (-0.071 to -0.052) and Recklessness (-0.087 to -0.068), consistent with partial mediation. Correspondingly, 27% and 22% of the effect of grade on COV and Recklessness, respectively, was mediated via EF. In the above mediation models, no significant EF by grade interactions were observed.

3.3. Relationship to school characteristics

BART-C measures (Table 2) did not demonstrate significant relationships to Overall Ratings of the schools; but schools with higher percentage of free or reduced school lunch (a proxy for poverty) had significantly lower BART-C Total and had higher Recklessness, significant at the trend level, $p = .055$). Children in schools with higher rates of Math and Reading proficiency were more likely to have better BART-C Total scores and to have lower Recklessness scores.

4. Discussion

4.1. Adaptive risk taking and self-regulation from K to 8th grade

BART-C extends the research done on the BART-Y [23] to a younger sample and offers data on the developmental trajectories of adaptive risk-taking, seen in our measures of Total score and Adjusted Puffs, and of self-regulation of emotion seen in reduced variation in number of puffs (COV) and in Recklessness. Although, there is a very strong linear relationship between grade and better Total score, higher Adjusted Puffs, and less variability (COV), it appears that Recklessness, evens out by grade 4 with a slight upward trend in middle school for the whole sample.

Table 1
Gender comparison of Standard Scores for each BART measure across grade.

Grade	BART Total			BART Adjusted			COV			Recklessness		
	Male LSmean	Female LSmean	StdErr	Male LSmean	Female LSmean	StdErr	Male LSmean	Female LSmean	StdErr	Male LSmean	Female LSmean	StdErr
K-garten	92.3	92.2	1.29	97.7	98.3	1.37	108.3	104.0	1.33	105.4	106.2	1.14
Grade 1	94.0	97.1	0.93	98.0	100.3	1.12	105.5	100.8	0.98	104.2	103.3	0.84
Grade 2	98.5	97.4	0.77	99.7	97.9	0.91	102.2	101.3	0.83	101.3	100.6	0.70
Grade 3	96.4	99.1	0.70	97.7	98.0	0.76	103.2	100.4	0.75	101.4	99.0	0.64
Grade 4	99.3	100.5	0.71	99.3	99.2	0.80	101.5	100.1	0.76	100.2	98.8	0.64
Grade 5	101.6	101.8	0.77	102.2	100.4	0.87	100.5	99.8	0.81	100.8	98.8	0.70
Grade 6	102.1	101.4	0.88	103.0	100.5	1.06	99.9	98.5	0.92	101.0	99.3	0.79
Grade 7	101.8	102.1	0.94	104.3	101.6	1.23	0.061	96.7	1.25	0.130	99.6	0.85
Grade 8	104.8	104.3	1.08	106.8	101.6	1.31	0.001	95.4	1.15	0.118	97.4	0.98

Note: Values are least-square means (LSmean) and standard errors (StdErr) estimated from each mixed model. P-values are based on post-hoc linear contrasts estimated from each mixed model.

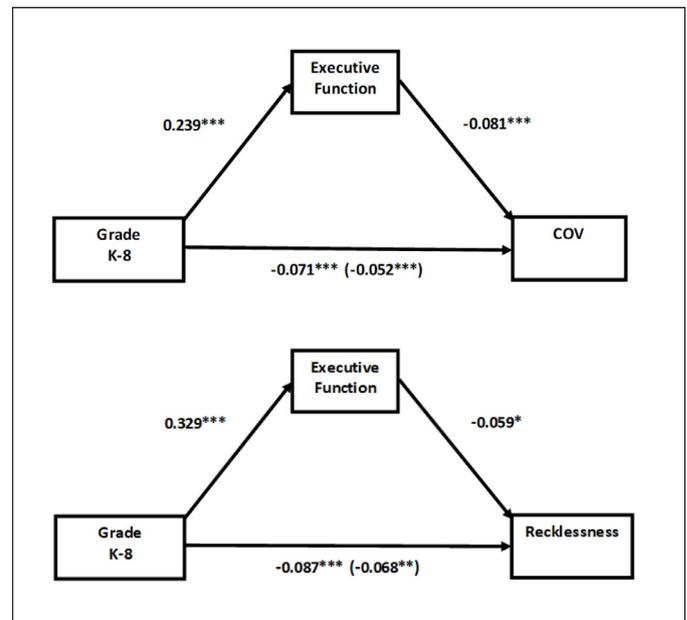


Fig. 4. Regression coefficients for the relation between Grade and COV, and Grade and Recklessness as mediated by Executive Function. COV: Total effect = -0.071 ; Direct effect = -0.052 ; Indirect effect = -0.019 ; % mediation = 27%. Note: the regression coefficient for Grade and COV, controlling for EF is in parentheses; *** $p < .0001$, ** $p < .001$, * $p < .01$. Recklessness: Total effect = -0.087 ; Direct effect = -0.068 ; Indirect effect = -0.019 ; % mediation = 22%. Note: the regression coefficient for Grade and Recklessness, controlling for EF is in parentheses; *** $p < .0001$, ** $p < .001$, * $p < .01$.

4.2. Gender differences

Gender differences were not significant as reported for BART-Y Adjusted Puffs in their sample of 9th through 12th graders (Mean age = 14.8, SD = 1.5; Percent Female X Adjusted Puffs, $r = 0.10$ [36]), but were significant in our 8th grade sample on Adjusted Puffs. Of potential significance is that gender differences on Recklessness increased and became significant for our sample in middle school (6th-8th), suggesting that girls needed to take less risks (Adjusted Puffs) to receive the same or better reward (Total score). We think that this finding may be indicating that early adolescent girls “maintain their cool” a little better than their male counter-parts in situations demanding “hot cognition”, and we encourage other investigators to adopt this potentially useful measure to explore this further.

4.3. Executive functions and BART-C measures

Although there is a clear developmental progression for our BART-C measures, there are also considerable individual differences. Using the mediation model suggested by Blair et al. [24], we found that EF was correlated with BART-C response variability (COV) and Recklessness and partially mediated the relationship between grade and those measures. Blair et al. found that their working memory measure but not a Global EF measure best mediated age and COV for their BART data, collected from adolescents (Mean age = 17.9, SD = 3.6). Our mediation analysis using COV with much younger children supports their finding, although the effects are somewhat weaker. We also found that EF partially mediated the effect of grade on Recklessness (K-4th grade), though again effects are small. Our findings suggest that while the development of adaptive risk-taking relates to the development of EF at the individual level, BART-C as a measure of “hot cognition” may be able to explain unique variance in how children function adaptively in

Table 2
Correlations of BART-C measures with School Characteristics (Ns vary by Niche reported characteristics).

		Overall Rating	Free or Reduced Lunch Percent	Reading	Math
BART-C Total	Pearson Correlation	.055	−0.280	.262	.276
	Sig. (2-tailed)	.694	.030	.047	.036
	N	54	60	58	58
BART-C Adjusted	Pearson Correlation	−0.068	−0.039	−0.237	−0.197
	Sig. (2-tailed)	.623	.766	.073	.137
	N	54	60	58	58
COV	Pearson Correlation	−0.155	.124	−0.004	−0.043
	Sig. (2-tailed)	.262	.346	.979	.746
	N	54	60	58	58
Recklessness	Pearson Correlation	−0.127	.249	−0.508	−0.481
	Sig. (2-tailed)	.361	.055	.000	.000
	N	54	60	58	58

their environment.

4.4. BART-C measures and school characteristics

We had no personally identifying information in our study, but we could access information about the schools our children attended. We correlated BART-C scores with school characteristics, including Overall Rating, Percent Free or Reduced Lunch and rates of Math and Reading proficiency according to Niche. Although there were no significant findings for Overall Ratings, children in schools with higher percentage of free or reduced school lunch (a proxy for poverty) had significantly lower Total score and had higher Recklessness, while children in schools with higher rates of Math and Reading proficiency had higher Total score and lower Recklessness. A previous study [37] using BART-Y with 224 young adolescents (Mean age = 11.0, SD = 0.9 years) failed to find significant relationships at the individual level between Family Income and BART-Y performance, so the association we found at the school level may not apply at the individual level. There is no previous report regarding BART scores and academic proficiency, although other measures of self-regulation have predicted better academic performance [38]. Interpretation of this finding with Recklessness should await individual data collection.

4.5. Limitations and future directions

In this report, we have described the first use of BART-C, a “hot cognition” measure in children K-8th grade. We believe that this child adaptation of the BART-Y is faithful to the original, but we recognize that exposing the child within the first 5 trials to an explosion after 2 puffs may have been more challenging to the child’s self-regulation than the conventional task. Thus, it remains unknown to what extent this study has continuity with previous studies using the original BART and the BART-Y. Study results support developmental trajectories for the BART-C measures, but findings are based on cross-sectional data based on grade rather than age, so any conclusive judgments about development of adaptive risk-taking must await age-related longitudinal studies.

There is also a great deal of individual difference within each grade. Our correlations with EF are modest and suggest that a great deal of variance in adaptive risk-taking cannot be understood by a child’s EF. Similarly, the relationship between grade and COV and Recklessness is only partially mediated by EF, further suggesting that there may be many other influences on the development of adaptive risk-taking not included in this model. We were working with a de-identified data set, so turned to the school characteristics as a proxy for some environmental factors that might play a role in children’s development of “hot cognition”. While we found some suggestive relationships with Recklessness, much more research is needed to explore individual differences and the factors that influence healthy development.

We recognize that COV is a relatively new measure in the BART

literature and that our Recklessness measure is original to this report. We are particularly encouraged by the Recklessness findings across the developmental span K to 8th grade, and the gender differences in older grades, and we hope that other investigators will explore the merits of the COV and Recklessness measure. We recognize that only repeated use of COV and Recklessness measures by other investigators will determine whether these are reliable measures of maladaptive risk-taking. This is the largest sample of children K-8th grade to have an assessment related to adaptive risk-taking. We hope that these findings will encourage use of BART-C, as a way to measure “hot cognition” in children.

Financial support

Support to Drs Bell, Imal and Wexler comes in part from NSF (Project #1,565,310) Cyber Human Systems Large: Collaborative Research: Computational Science for Improving Assessment of Executive Function in Children (Bell, PI).

Support to Dr. Bell comes in part from VA Rehabilitation Research & Development Senior Research Career Scientist Award (1 IK6 RX002458-01)

Ethical statement

This study used only de-identified archival data. It was determined to meet the requirement under Category 4 for exemption from IRB review.

Declaration of Competing Interest

Bell is an investor in C8 Sciences, a Yale University sponsored Start-up company, which collected the data for this study.

Imal is a post-doctoral fellow partially funded by C8 Sciences.

Pittman serves as statistician independent of C8 Sciences and has no potential conflict of interest.

Jin is on a NSF Research Experience for Undergraduates paid work experience at C8 Science.

Wexler is an investor in C8 Sciences and is on the Board of Directors.

We do not believe that our involvement in C8 Sciences influenced the findings. Most analyses were performed by Pittman, who is independent of C8 Sciences.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.tine.2019.100120](https://doi.org/10.1016/j.tine.2019.100120).

References

- [1] A. Schore, Affect regulation and the origin of the self: The neurobiology of

- emotional development, 1996. doi:10.1097/00004583-199611000-00028.
- [2] W. Hofmann, B.J. Schmeichel, A.D. Baddeley, Executive functions and self-regulation, *Trends Cogn. Sci.* (2012), <https://doi.org/10.1016/j.tics.2012.01.006>.
- [3] M. Sperduti, D. Makowski, M. Arcangeli, P. Wantzen, T. Zalla, S. Lemaire, J. Dokić, J. Pelletier, P. Piolino, The distinctive role of executive functions in implicit emotion regulation, *Acta Psychol. (Amst.)* (2017), <https://doi.org/10.1016/j.actpsy.2016.12.001>.
- [4] C. Blair, C.C. Raver, Closing the achievement gap through modification of neuro-cognitive and neuroendocrine function: Results from a cluster randomized controlled trial of an innovative approach to the education of children in kindergarten, *PLoS ONE* (2014), <https://doi.org/10.1371/journal.pone.0112393>.
- [5] L. Nota, S. Soresi, B.J. Zimmerman, Self-regulation and academic achievement and resilience: A longitudinal study, *Int. J. Educ. Res.* (2004), <https://doi.org/10.1016/j.ijer.2005.07.001>.
- [6] M.M. McClelland, A.C. Acock, F.J. Morrison, The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school, *Early Child. Res. Q.* (2006), <https://doi.org/10.1016/j.ecresq.2006.09.003>.
- [7] S.A. Schmitt, M.M. McClelland, S.L. Tominey, A.C. Acock, Strengthening school readiness for head start children: Evaluation of a self-regulation intervention, *Early Child. Res. Q.* (2015), <https://doi.org/10.1016/j.ecresq.2014.08.001>.
- [8] K. Kinniburgh, K. Jentoft Kinniburgh, M. Blaustein, J. Spinazzola, Attachment, self-regulation, and competency: A comprehensive intervention framework for children with complex trauma, *Psychiatr. Ann.* (2005).
- [9] T.J. Dishion, A. Connell, Adolescents' resilience as a self-regulatory process: Promising themes for linking intervention with developmental science, *Ann. N. Y. Acad. Sci.* (2006), <https://doi.org/10.1196/annals.1376.012>.
- [10] S.P. Hinshaw, Process, mechanism, and explanation related to externalizing behavior in developmental psychopathology, *J. Abnorm. Child Psychol.* (2002), <https://doi.org/10.1023/A:1019808712868>.
- [11] E. Kuntsche, R. Knibbe, R. Engels, G. Gmel, Bullying and fighting among adolescents - Do drinking motives and alcohol use matter? *Addict. Behav.* (2007), <https://doi.org/10.1016/j.addbeh.2007.07.003>.
- [12] P.A. Wyman, P.A. Gaudieri, K. Schmeelk-Cone, W. Cross, C.H. Brown, L. Sworts, J. West, K.C. Burke, J. Nathan, Emotional triggers and psychopathology associated with suicidal ideation in urban children with elevated aggressive-disruptive behavior, *J. Abnorm. Child Psychol.* (2009), <https://doi.org/10.1007/s10802-009-9330-4>.
- [13] A. Diamond, K. Lee, Interventions shown to aid executive function development in children 4 to 12 years old, *Science* (80) (2011), <https://doi.org/10.1126/science.1204529>.
- [14] P.A. Wyman, W. Cross, C.H. Brown, Q. Yu, X. Tu, S. Eberly, Intervention to strengthen emotional self-regulation in children with emerging mental health problems: Proximal impact on school behavior, *J. Abnorm. Child Psychol.* (2010), <https://doi.org/10.1007/s10802-010-9398-x>.
- [15] A.F.T. Arnsten, Stress signalling pathways that impair prefrontal cortex structure and function, *Nat. Rev. Neurosci.* (2009), <https://doi.org/10.1038/nrn2648>.
- [16] B.P. Ramos, A.F.T. Arnsten, Adrenergic pharmacology and cognition: Focus on the prefrontal cortex, *Pharmacol. Ther.* (2007), <https://doi.org/10.1016/j.pharmthera.2006.11.006>.
- [17] C. Blair, D. Granger, R.P. Razza, Cortisol reactivity is positively related to executive function in preschool children attending head start, *Child Dev.* (2005), <https://doi.org/10.1111/j.1467-8624.2005.00863.x>.
- [18] M.B. Bronson, T. Tivnan, P.S. Seppanen, Relations between teacher and classroom activity variables and the classroom behaviors of prekindergarten children in chapter 1 funded programs, *J. Appl. Dev. Psychol.* (1995), [https://doi.org/10.1016/0193-3973\(95\)90035-7](https://doi.org/10.1016/0193-3973(95)90035-7).
- [19] G. a Gioia, P.K. Isquith, S.C. Guy, L. Kenworthy, Behavior rating inventory of executive function - Ensignant, *Child Neuropsychol.* (2000), <https://doi.org/10.1076/chin.6.3.235.3152>.
- [20] T.M. Achenbach, L. Rescorla, Manual for the ASEBA school-age forms & profiles: Child behavior checklist for ages 6-18, teacher's report form, youth self-report, An Integr. Syst. Multi-Informant Assess. (2001).
- [21] M.M. McClelland, C.E. Cameron, R. Duncan, R.P. Bowles, A.C. Acock, A. Miao, M.E. Pratt, Predictors of early growth in academic achievement: The head-toes-knees-shoulders task, *Front. Psychol.* (2014), <https://doi.org/10.3389/fpsyg.2014.00599>.
- [22] S.E. Petersen, M.I. Posner, The attention system of the human brain: 20 years after, *Annu. Rev. Neurosci.* (2012), <https://doi.org/10.1146/annurev-neuro-062111-150525>.
- [23] C.W. Lejuez, J.B. Richards, J.P. Read, C.W. Kahler, S.E. Ramsey, G.L. Stuart, D.R. Strong, R.A. Brown, Evaluation of a behavioral measure of risk taking: The balloon analogue risk task (BART), *J. Exp. Psychol. Appl.* (2002), <https://doi.org/10.1037/1076-898X.8.2.75>.
- [24] M.A. Blair, A. Moyett, A.A. Bato, P. DeRosse, K.H. Karlsgodt, The role of executive function in adolescent adaptive risk-taking on the balloon analogue risk task, *Dev. Neuropsychol.* (2018), <https://doi.org/10.1080/87565641.2018.1510500>.
- [25] C.W. Lejuez, W.M. Aklin, M.J. Zvolensky, C.M. Pedulla, Evaluation of the balloon analogue risk task (BART) as a predictor of adolescent real-world risk-taking behaviours, *J. Adolesc.* (2003), [https://doi.org/10.1016/S0140-1971\(03\)00036-8](https://doi.org/10.1016/S0140-1971(03)00036-8).
- [26] C.W. Lejuez, W.M. Aklin, H.A. Jones, D.R. Strong, J.B. Richards, C.W. Kahler, J.P. Read, The balloon analogue risk task (BART) differentiates smokers and non-smokers, *Exp. Clin. Psychopharmacol* (2003), <https://doi.org/10.1037/1064-1297.11.1.26>.
- [27] K.L. Hanson, R.E. Thayer, S.F. Tapert, Adolescent marijuana users have elevated risk-taking on the balloon analog risk task, *J. Psychopharmacol.* (2014), <https://doi.org/10.1177/0269881114550352>.
- [28] M.A. Bornovalova, S.B. Daughters, G.D. Hernandez, J.B. Richards, C.W. Lejuez, Differences in impulsivity and risk-taking propensity between primary users of crack cocaine and primary users of heroin in a residential substance-use program, *Exp. Clin. Psychopharmacol* (2005), <https://doi.org/10.1037/1064-1297.13.4.311>.
- [29] T.J. Crowley, K.M. Raymond, S.K. Mikulich-Gilbertson, L.L. Thompson, C.W. Lejuez, A risk-taking "set" in a novel task among adolescents with serious conduct and substance problems, *J. Am. Acad. Child Adolesc. Psychiatry.* (2006), <https://doi.org/10.1097/01.chi.0000188893.60551.31>.
- [30] W.M. Aklin, C.W. Lejuez, M.J. Zvolensky, C.W. Kahler, M. Gwadz, Evaluation of behavioral measures of risk taking propensity with inner city adolescents, *Behav. Res. Ther.* (2005), <https://doi.org/10.1016/j.brat.2003.12.007>.
- [31] M.D. Bell, H.B. Laws, I.B. Petrakis, A randomized controlled trial of cognitive remediation and work therapy in the early phase of substance use disorder recovery for older veterans: Neurocognitive and substance use outcomes, *Psychiatr. Rehabil. J.* (2017), <https://doi.org/10.1037/prj0000211>.
- [32] K.S. DeMartini, R.F. Leeman, W.R. Corbin, B.A. Toll, L.M. Fucito, C.W. Lejuez, S.S. O'Malley, A new look at risk-taking: Using a translational approach to examine risk-taking behavior on the balloon analogue risk task, *Exp. Clin. Psychopharmacol.* (2014), <https://doi.org/10.1037/a0037421>.
- [33] P.J. Diggle, K.Y. Liang, S.L. Zeger, *Analysis of Longitudinal Data*, Oxford University Press, New York, 1994.
- [34] R.M. Baron, D.A. Kenny, The moderator-mediator variable distinction in social psychological research. conceptual, strategic, and statistical considerations, *J. Pers. Soc. Psychol.* (1986), <https://doi.org/10.1037/0022-3514.51.6.1173>.
- [35] M.E. Sobel, S. Leinhardt, Asymptotic confidence intervals for indirect effects in structural equation models, *Sociol. Methodol.* (1982) (1982), <https://doi.org/10.2307/270723>.
- [36] C.W. Lejuez, W. Aklin, S. Daughters, M. Zvolensky, C. Kahler, M. Gwadz, Reliability and validity of the youth version of the balloon analogue risk task (BART-Y) in the assessment of risk-taking behavior among inner-city adolescents, *J. Clin. Child Adolesc. Psychol.* (2007), <https://doi.org/10.1080/15374410709336573>.
- [37] A. Collado, C.M. Risco, A.N. Banducci, K.W. Chen, L. MacPherson, C.W. Lejuez, Racial differences in risk-taking propensity on the youth version of the balloon analogue risk task (BART-Y), *J. Early Adolesc.* (2017), <https://doi.org/10.1177/0272431615596429>.
- [38] M.M. McClelland, C.E. Cameron, Self-regulation and academic achievement in elementary school children, *New Dir. Child Adolesc. Dev.* (2011), <https://doi.org/10.1002/cd.302>.