



The Effectiveness of a Teacher-Delivered Mindfulness-Based Curriculum on Adolescent Social-Emotional and Executive Functioning

Jennifer L. Frank¹ · Patricia C. Broderick¹ · Yoonkyung Oh¹ · Joy Mitra¹ · Kimberly Kohler¹ · Deborah L. Schussler¹ · Charles Geier¹ · Robert W. Roeser¹ · Elaine Berrena¹ · Julia Mahfouz¹ · Joseph Levitan¹ · Mark T. Greenberg¹

Accepted: 15 January 2021 / Published online: 16 February 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC part of Springer Nature 2021

Abstract

Objectives The purpose of this study was to assess the effectiveness of a manualized mindfulness-based program for adolescents, Learning to Breathe (L2B), on indicators of adolescent social-emotional well-being, mental health, substance use, and executive function.

Methods Participants included 251 high school students attending an urban school district who were randomly assigned to required health education classes that offered L2B or the business-as-usual health curriculum.

Results No direct effects on self-report measures were found. Students exposed to L2B demonstrated significant improvements on executive functioning measures related to susceptibility to cognitive interference and working memory. Subsequent tests of moderation revealed beneficial program effects within the treatment group were dependent on rates of practice, and those that used L2B strategies regularly showed small-to-moderate improvements on indices of emotional awareness, emotional clarity, impulse control, social connectedness, mind-wandering, substance use, perceived stress, and self-compassion relative to controls.

Conclusions This study provides mixed support regarding the potential effectiveness of a universal mindfulness program for high school students. The absence of direct effects on self-report measures implies that simply exposing adolescents to a mindfulness curriculum within the context of typical instruction, in the absence of supports for implementation, is unlikely to impact youth socio-emotional well-being or behavior. However, changes in EF favoring the intervention group suggest that possible benefits on tasks related to susceptibility, cognitive interference, and selective attention are possible. Tests of moderation revealed dosage effects, and students who adopt mindfulness practices can indeed benefit on multiple fronts.

Keywords L2B · Mindfulness · Adolescence · Learning to Breathe

Adolescence is a developmental period characterized by both opportunity and risk. The cognitive, physical, social, and emotional advances observed in this decade are important foundations for adult responsibility and independence. At the same time, emotional and behavioral problems increase and represent the primary cause of adolescent mortality and morbidity. Notably, the dramatic neurological changes that occur create a window of vulnerability and are associated with the onset of

major mental illnesses such as anxiety, depression, and substance abuse (Dahl 2004; Greenberg and Lippold 2013). Experience shapes adolescents' sensitive neural systems and may affect long-term mental health (Casey 2015).

Mindfulness-based programs (MBPs) show some promise for supporting healthy adolescent development and well-being. Defined as programs that train attentional skills in a certain way: with intention, present-moment focus, and acceptance (Kabat-Zinn 1990), MBPs teach and practice an approach to experience that is open, curious, and non-reactive.

Systematic reviews have concluded that MBPs are feasible within school settings and probably efficacious for anxiety, depression, and stress reduction (Black and Slavich 2016; Felver et al. 2016). Maynard et al. (2017) concluded that there are mixed effects of MBPs in schools, with some indication that MBPs can improve cognitive and socio-emotional

✉ Jennifer L. Frank
jfrank@psu.edu

¹ Department of Educational Psychology, Counseling, and Special Education, The Pennsylvania State University, 308 Cedar Building, University Park, PA 16802-3108, USA

outcomes but no support for improvement in behavior or academic achievement. They note that despite the growing support of MBPs for adults, youth may not benefit in the same ways or to the same extent as adults. Another meta-analysis of school-based studies (Carsley et al. 2018) found greater effects for older vs. younger adolescents and overall effects for well-being were found only for MBPs delivered by trained teachers rather than outside facilitators. However, there is a dearth of research on effects of MBPs implemented by actual classroom teachers in real-world settings.

Researchers have identified several competencies necessary for the effective delivery of MBPs including adequate coverage, pacing and organization of curriculum, interpersonal relational skills, skillful guiding of formal meditation practices, effective interactive inquiry, group dialogue and didactic teaching, and management of the learning environment. One school of thought suggests that effective mindfulness teachers share a common commitment to mindfulness through regular daily practice, which over time, manifests itself as “embodiment” or the authentic expression of the skills and dispositions necessary for effective mindfulness instruction (Crane et al. 2016; Roeser 2016). A key challenge to the portability and sustainability of MBPs is whether classroom teachers can be adequately trained to deliver such curricula in the context of typical classroom settings and constraints.

Despite the challenges involved, adolescents may be particularly receptive to the benefits of mindfulness training. Current evidence suggests that the plasticity of the adolescent brain makes this a sensitive period of development, maximally receptive to environmental inputs (Blakemore and Mills 2014). The prefrontal cortex (PFC), which integrates emotional and contextual information and has important regulatory functions, is being fine-tuned, with neuromaturation continuing into adulthood (Caballero et al. 2016). Subcortical brain regions related to mood, threat assessment (Davidson 2000), and reward-seeking (e.g., nucleus accumbens and ventral striatum; Gottfried 2011) also continue to develop and increase in reactivity. One frequently cited implication of these protracted developmental processes is that the executive and regulatory capacities of the adolescent brain have great difficulty overriding strong emotionality and reward-seeking, a model of adolescent brain development known as a “developmental mismatch” or “maturational imbalance” (Casey 2015; Somerville et al. 2010).

Experimental evidence shows that, compared to children and adults, adolescents display greater amygdala arousal (Hare et al. 2008), report more daily experience of negative affect from age 10 to 18 (Larson et al. 2002), are more inclined to maintain negative affect (Riediger et al. 2009), and are more likely to act impulsively when they perceive cues as threatening (Dreyfuss et al. 2014). Adolescents, compared to other age groups, are also less effective at maintaining attention on task in the presence of emotional stimuli, particularly when the

stimuli are distressing (Cohen-Gilbert and Thomas 2013). The executive functions (EFs) of shifting and monitoring of attention, and the planning, initiating, and carrying out of goal-directed behavior are still developing (Casey 2015). Consequently, lower levels of top-down regulation (EFs) in the context of heightened emotionality create an imbalance that can predispose adolescents to risky behavior and maladaptive choices.

Emotion regulation (ER), defined as processes used to moderate affective experiences in order to meet situational demands, include modulating arousal, inhibiting automatic responses, persisting during stressful activities, and delaying gratification (Gross and Thompson 2007). Not surprisingly, difficulties in ER represent a core feature of many emotional and behavioral symptoms and disorders that emerge in adolescence (Powers and Casey 2015).

Despite the fact that EF and ER are considered separate constructs, it is important to recognize that brain circuitry related to both emotion and executive processing is completely intertwined and that cognitive and affective functions are not dissociable. Poorly controlled emotions can directly impair information processing and other goal-directed EFs (Best et al. 2009) and degrade academic performance (Rothbart et al. 2007). Similarly, poor cognitive control limits the power to regulate emotions and may increase rumination increasing risk for internalizing disorders in adolescents (Snyder and Hankin 2016).

Cognitive control is central to theories of ER. Strategies such as selecting and modifying situations, anticipating and planning outcomes, attention management, and cognitive reappraisal rely heavily on EF capacities. While top-down regulatory strategies are typically available to adults, they may be far less so for adolescents. Adolescents’ heightened emotionality and contextual sensitivity may compromise EFs and make strategy selection and implementation more difficult. Furthermore, adolescents’ ability to identify and differentiate negative emotions (angry, disgusted, sad, scared, and upset) is at the low point of a U-shaped developmental curve (Nook et al. 2018). Therefore, intervention approaches for adolescents need to take the normative developmental conditions of high arousal, greater negative affect, underdeveloped EF skills, and context-sensitivity into account.

The fundamental practice of mindfulness involves paying attention, either to a specific focus of attention or to the array of phenomenal experience, with an attitude of curiosity, nonjudgment, and acceptance (Kabat-Zinn 1990). Awareness and non-reactive experiencing of thoughts, emotions, and sensations, without suppression or avoidance, may minimize the risk of behavioral consequences like rumination or acting out (Chambers et al. 2009). Mindfulness practice exercises both executive (maintaining and shifting attention, inhibiting distractions) and affective (strengthening limbic and prefrontal regulatory circuitry) systems (Tang et al. 2015).

Iterative reprocessing of information, as is done by reflectively returning attention again and again to a certain object of focus, promotes cognitive control and flexibility (Zelazo 2015). Therefore, mindfulness practice may be a promising means of scaffolding regulation and reflection during a period of rapid neurobiological change.

Learning to BREATHE (L2B) is an MBP for adolescents designed to strengthen ER and EF (Broderick 2013; see Fig. 1). The core curriculum provides developmentally adapted training in several core practices of mindfulness-based stress reduction (MBSR; Kabat-Zinn 1990) including body scan, awareness of thoughts and feelings, mindful movement, and loving-kindness/compassion practices.

L2B outcomes on student well-being and learning have been explored in several studies. Early quasi-experimental studies showed reduced negative affect and increased well-being in 12th grade private school girls following L2B (Broderick and Metz 2009) and reduced stress, somatic complaints, and higher perceived emotion regulation efficacy compared to controls in public high school students (Metz et al. 2013). Three pilot randomized controlled studies have expanded this work. Ethnically diverse alternative high school students assigned to L2B, compared to a substance abuse prevention class (total sample was 27), demonstrated lower levels of depression but no significant effects on anxiety, mindfulness, and perceived stress compared to controls (Bluth et al. 2016). Fung et al. (2019) found improvements on stress reduction, internalizing problems and emotion regulation for predominantly Asian-American and Latino-American 9th

graders with elevated mood symptoms after L2B and at 3-month follow-up. L2B participation did not lead to improvements in externalizing problems, attention problems, and expressive suppression. In a study involving at-risk high school students (Felder et al. 2018), levels of resilience in the L2B group were maintained over time compared to controls, but there were no intervention effects on self-reported problem behavior, school attendance, and quarterly academic grades.

Although the results of these studies are promising, most outcomes are based on self-report measures and L2B has yet to be examined in a large independent trial with classroom teachers as instructors. The goal of the present study was to assess the effectiveness of L2B, delivered by trained high school teachers, in authentic educational settings using a quasi-experimental control-group trial design. Furthermore, we utilized both student self-report measures and direct assessment of EF outcomes at post-test. We hypothesized that consistent with the L2B logic model, L2B would significantly reduce adolescents' perceived stress, decrease their symptoms of anxiety and depression, improve their EFs, and enhance their well-being.

Method

Participants

Informed consent was obtained from all participants. As shown in the CONSORT diagram (Fig. 2), letters of consent

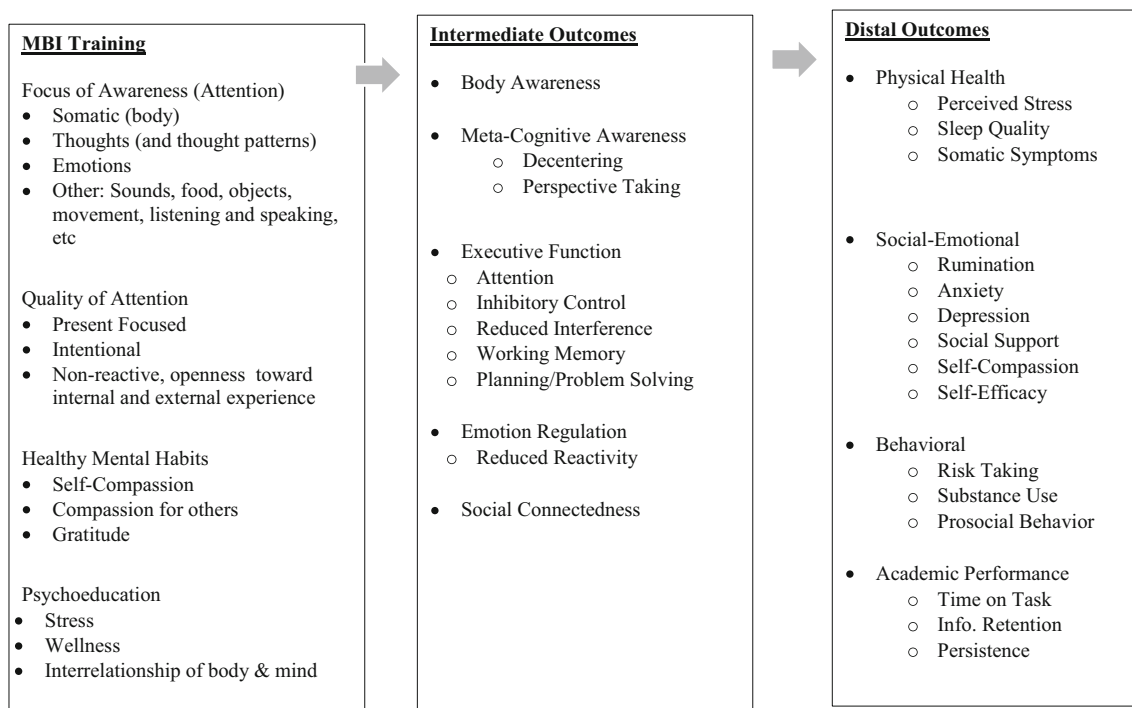
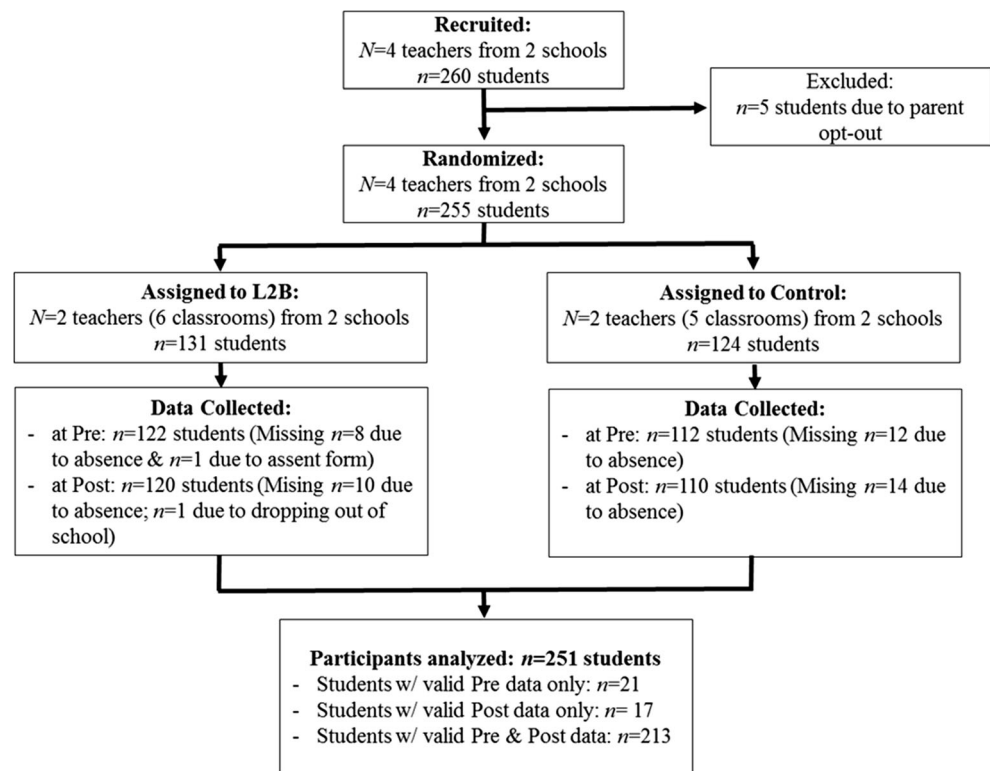


Fig. 1 Learning to Breathe (L2B) logic model

Fig. 2 CONSORT diagram of participant flow through the phases of the study



were sent to parents of 260 students recruited from the classrooms of four teachers (two intervention and two control). Teachers taught multiple independent course sections (intervention = 6 classes, and control = 5 classes). Our study design blocked on school, and therefore, there was one intervention teacher and one control teacher nested within each school. During recruitment, five parents opted-out of participation, and 255 students were randomized to L2B or business-as-usual control conditions. A total of 131 students were enrolled in six classrooms of the teachers assigned to implement L2B (intervention group), and 124 students enrolled in the five classrooms of the control teachers (control group). Of the 255 randomized students, four were absent at both pre- and post-data collection; thus, the final analytic sample was 251 students. The racial/ethnic composition was diverse: 50% White, 16% Black, 9% Hispanic, and 5.6% Asian. About 57% was male, and 23% received free lunch. The mean age was 16 years old. About 55% reside in two-biological-parent families, 19% with stepparents, and 19.5% with single parents.

Procedure

The study took place in two suburban high schools in the Northeast US. L2B was implemented during required health education classes in 11th grade. Two teachers volunteered to be trained in L2B and two other health teachers and their classrooms served as a business-as-usual control. After

blocking on school, students were randomly assigned to one of 11 classrooms by administrators (intervention classrooms $n = 6$, control classrooms $n = 5$) taught by these four health teachers. Students assigned to control were exposed to the approved high school health curriculum which included units related to mental health (e.g., dimensions of wellness, disorders), social health (health and abuse relationships, bullying), human sexuality (e.g., human reproduction, contraception, sexually transmitted infections), substance abuse (e.g., effects of alcohol, tobacco, and other drugs), and nutrition and fitness (e.g., principles of nutrition, eating disorders, and body image).

Intervention The L2B program is organized according to a scope and sequence of content units built around the “BREATHE” acronym (body, reflections, emotions, attention), in the current study, was delivered in 12 sessions (two sessions for each unit). Table 1 provides a brief overview of each unit theme and associated procedures, and Table 2 provides an overview of core intervention components including training and discussion topics, experiential activities, specific mindfulness-based practices, and homework practices. The first unit provides an introduction to mindfulness and practice in somatic awareness. The second unit includes activities to understand and identify automatic self-talk and using mindfulness-based strategies to approach it. The third unit explores how emotions affect thoughts and somatic sensations. The fourth session focuses on understanding stress

Table 1 Overview of L2B curriculum sessions and procedures

Class session	Letter	Theme	Procedures
1	B	Listen: Your body is trying to tell you something.	Introduction to mindfulness and mindlessness; practice in somatic awareness.
2	R	Reflections (thoughts) are just thoughts.	Discussion, activities, and practice designed to understand automatic self-talk and to approach it using mindfulness.
3	E	Surf the waves of your emotions.	Discussion, activities, and practice designed to understand how emotions affect thoughts and bodily sensations.
4	A	Attention to body, thoughts and feelings is good stress management.	Discussion, activities, and practice designed to understand stress and stress reactivity.
5	T	Take it as it is (non-judgment).	Discussion, activities, and practice designed to cultivate loving kindness
6	H	Practice healthy mind habits to reduce stress and increases inner strength.	Wrap-up of all sessions and discussion of how to use mindfulness in one's life.

and stress reactivity. The fifth unit focuses on practices designed to cultivate compassion towards the self and others through loving-kindness practice. The sixth unit includes a wrap-up of prior sessions and discussion of how to integrate mindfulness into one's daily life. Each L2B lesson follows a predictable format including a short introduction, activities for group participation and discussion to engage students in the lesson, and in-class mindfulness practice. Workbooks and CDs for home mindfulness practice are provided as a supplement. L2B content includes the core practices of MBSR including body scan, awareness of thoughts and feelings, mindful movement, and loving kindness practice developmentally adapted for adolescents.

Teacher Training Prior to training teachers in L2B, the program developer and on-site coach met with intervention-condition teachers for four weekly individual training sessions (6 h total) to orient teachers to the concept of mindfulness and help them to establish a personal practice prior to teaching L2B. The sessions included practice in mindfulness strategies, weekly practice assignments, and journal writing to promote a personal practice. Subsequently, teachers attended a 2-day training (14 h total), led by the program developer, on implementing L2B curriculum with fidelity in their classrooms. The L2B training included practice teaching each

lesson segment in small groups and sharing peer feedback, and discussion of the pedagogy of embodied mindful teaching. Materials were shared to assist teachers to recognize key transitions in each lesson and to balance use of time between activities, discussion, and mindfulness practice.

Coaching During L2B implementation, five weekly coaching calls (60 min) occurred during the 12-lesson sequence over a 6-week period. Prior to the weekly calls, coaches (program developer and second coach who had several years of prior experience in delivering mindfulness-based training and support to teachers in school settings) viewed lesson videotapes. Coaching sessions focused on L2B implementation, including lesson content and delivery, inquiry and experiential practices, interpersonal relational skills, guidance of mindfulness practices, managing and holding the space for student and teacher experiences, and embodying the qualities and characteristics associated with mindfulness (Crane et al. 2016). In addition to calls, coaches made one classroom visit with each teacher during the study.

Thus, teachers were given background practice in mindfulness, and received a workshop on implementation of L2B and on-site and phone coaching support during implementation. One intervention group teacher had prior experience and training in MBP strategies for adults and had a regular meditation practice (> once per week) that was established for over a year. The remaining intervention and control group teachers did not have prior experience or training in MBPs and did not meditate on a regular basis prior to starting the L2B program.

Fidelity Intervention fidelity (i.e., adherence) to manualized L2B lesson components was assessed by independent coders ($n = 7$) who were randomly assigned to code each lesson for fidelity. The intervention fidelity coding measure and manual were created by the program developer and included a point-by-point listing of all required actions and activities for each individual intervention lesson. All coders independently completed a partial day training on fidelity coding procedures led by the intervention developer and were required to achieve 80% or better reliability with master codes on training videos before coding study videos. All videos were coded by at least two coders. Average inter-coder agreement was 92% and disagreements were resolved by a third independent coder (study PI) in consultation with the program developer. The overall fidelity of implementation across sessions was 78.60% across teachers.

Measures

Measures were selected based on previous research and the L2B logic model which hypothesizes that L2B has direct immediate effects on measures of mindfulness, EF and ER skills, and student mental and physical health.

Table 2 L2B core intervention components

Content & discussion topics	Experiential activities	Mindfulness practice	Homework practice
1. Introduction to mindfulness and the concept of empowerment as balance.	1. Group discussion of mindless and mindful times in students' lives.	1. Mindful listening or mindful eating.	1. Journal reflections in student workbook.
2. Exploring how automatic thoughts affect behavior, mood and ability to concentrate.	2. Recognizing the chattering mind and automatic thoughts.	2. Mindfulness of the body.	2. Short body-scan meditation (CD).
3. Discussion of emotions as friends or foes.	3. Listening to the emotions evoked by music and recognizing their effects.	3. Mindfulness of thoughts.	3. Short meditation on thoughts (CD).
4. Exploring bodily awareness of emotions.	4. Drawing the effects of stress on the body.	4. Mindfulness of emotions.	4. Short meditation on emotions (CD).
5. Didactic information about the impact of stress on the body, thoughts, and feelings.	5. Identifying your place on the chronic stress curve.	5. Mindful movement.	5. Short
6. Distinctions between acute and chronic stress (daily hassles).	6. Activity to list ways we practice meanness (non-acceptance) of self and experience.	6. Loving-kindness meditation.	
7. Using mindful awareness to recognize and manage distress as it manifests in the body, thoughts, or feelings.	7. Listing top-ten stressors activity.	7. Guided imagery meditation on balance and resilience.	
8. Exploring how we practice meanness (non-acceptance) to self and experience.			
9. Exploring the benefits of self-compassion and acceptance of experience.			
10. Information about ways to generalize the practice of mindfulness. Loving-kindness practice (CD).			

Students completed a 30-min online battery of self-report measures and a 30-min battery of computer-administered EF measures. Pre-test assessments were administered 1 week prior to the intervention in the fall semester. Post-test assessments were completed 1 week after the program ended at the end of the fall semester. Scale scores were computed for respondents providing at least 80% of the items. Cronbach alphas were used to assess measure reliability at baseline.

Mindfulness Mindfulness was measured using the *Child and Adolescent Mindfulness Measure* (CAMM; Greco et al. 2011). The CAMM is a 10-item measure to assess mindfulness skills in children and adolescents. The CAMM asks respondents to rate on a 5-point scale (1 = never true to 5 = always true) the frequency with which they experience

feelings, thoughts, and behaviors that reflect a lack of mindfulness (e.g., *I think about things that happened in the past instead of thinking about things that are happening right now*). Items were summed and higher scores represent greater mindfulness ($\alpha = .89$).

Self-Compassion Self-compassion was measured with the *Self-Compassion Scale - Short Form* (SCS-SF; Raes et al. 2011). The SCS-SF is a 12-item measure designed to assess the ability to be compassionate to oneself. The SCS-SF asks respondents to rate on a 5-point scale (1 = almost never to 5 = almost always) the frequency with which they behave in a manner that reflects self-compassion (e.g., *When I'm going through a very hard time, I give myself the caring and tenderness I need.*). The SCS-SF includes six dimensions of self-compassion, including self-

kindness, self-judgment, common humanity, isolation, mindfulness, and over-identification. Items are summed to create a composite in which higher scores reflect greater self-compassion ($\alpha = .77$).

Emotion Regulation Emotion regulation was measured with the *Difficulties in Emotion Regulation Scale* (DERS; Gratz and Roemer 2004). The DERS is a 41-item questionnaire designed to assess five dimensions of emotion regulation difficulties: emotional awareness, emotional clarity, impulse control, access to emotion regulation strategies, and engaging in goal-directed behaviors. The DERS asks respondents to rate on a 5-point scale (1 = almost never to 5 = almost always) the frequency with which they experience emotion regulation difficulties (e.g., *When I'm upset, I have difficulty getting work done*). Items are summed to form subscales and an overall composite, with higher scores reflecting greater difficulties in emotion regulation. Coefficient alpha for the composite scale was .86, and subscales ranged from .50 to .87. The impulse control difficulties ($\alpha = .51$), and limited access to emotion regulation strategies ($\alpha = .65$) subscales had internal consistencies less than .80 and should be interpreted with caution.

Depression Depressive symptoms were measured with the *Patient Health Questionnaire* (PHQ-8; Kroenke et al. 2009). The PHQ-8 is an 8-item measure which asks respondents to rate on a 5-point scale (0 = not at all to 4 = nearly every day) the frequency with which they experience depression-related symptoms (e.g., *Little interest or pleasure in doing things*). Items were summed, with higher scores indicating higher levels of depression ($\alpha = .83$).

Anxiety Anxiety-related symptoms were measured using the Generalized Anxiety Disorder Scale (GAD-7; Spitzer et al. 2006). The GAD-7 asks respondents to report the frequency with which they experience generalized anxiety symptoms (e.g., *Trouble relaxing*) on a 4-point scale (0 = not at all to 4 = nearly every day). Items were summed, with higher scores indicating higher levels of anxiety ($\alpha = .89$).

Rumination Rumination was measured with the Rumination subscale of the *Rumination and Reflection Questionnaire* (RRQ; Trapnell and Campbell 1999). The RRQ rumination items are rated on a 5-point Likert scale (1 = strongly disagree to 5 = strongly agree). A priori, 9 of 12 rumination items (e.g., *dwell over things, thinking back over embarrassing moments, reevaluating*) were chosen that were most closely aligned to the L2B logic model. Items were summed with higher scores indicating higher levels of rumination ($\alpha = .85$).

Stress Stress was measured with two subscales of the *Adolescent Stress Questionnaire* (ASQ; Caballero et al.

2016) which measures the level of distress experienced in relation to common sources of adolescent stress. We used the subset of the ASQ items assessing two dimensions of adolescent stress; *stress of school performance* (6 items) and *stress of peer pressure* (5 items). The items are rated on a 5-point Likert scale (1 = not at all stressful to 5 very stressful) and summed for each dimension ($\alpha > .90$ for both scales).

Somatization Somatization was measured with the short-form of the *Children's Somatization Inventory* (CSI; Walker et al. 2008) which consists of 13 items that assess the frequency of stress-related physical symptoms (e.g., headaches, nausea) on a scale from 1 (very slightly) to 5 (almost always). Items were summed to create a total scale score ($\alpha = .92$).

Sleep The *Adolescent Sleep-Wake Scale* (ASWS; LeBourgeois et al. 2005) assesses overall sleep quality including falling asleep (3 items), maintaining sleep (6 items), and reinstating sleep (3 items) during the last month using a six-point, Likert-type scale (1 = always to 6 = never), higher scores representing better sleep quality ($\alpha = .86$).

Social Connectedness The *Social Connectedness Scale-Revised* (SCC-R; Lee et al. 2001) measures the extent to which individuals feel socially connected to individuals and social groups. Items are rated on a six-point scale (1 = strongly disagree to 6 = strongly agree) with higher scores representing greater social connectedness. From the original 20-item scale, we utilized 12 items with the highest factor loadings to form a single composite score ($\alpha = .90$).

Mind Wandering The *Mind Wandering Questionnaire* (MWQ; Mrazek et al. 2013) is a five-item scale measuring the frequency of interruption of task focus by unrelated thoughts. Items are measured on a six-point scale ($\alpha = .91$).

Growth Mindset Growth mindset was measured with the *Implicit Theories of Intelligence Scale for Children* (IT; Dweck 1999). The IT assesses individual beliefs regarding the fixed vs. malleable nature of their intelligence (e.g., *You can always greatly change how intelligent you are*). Items are rated on a six-point scale with higher scores indicating endorsement of an incremental theory of intelligence ($\alpha = .80$).

Substance Use Substance use was assessed using the *Substance Initiation Index* (Spath et al. 2007). *Gateway substance use* was measured by summing the responses to six binary items asking about lifetime experience of having a drink of alcohol, drinking more than a few sips of alcohol, having been drunk, smoking a cigarette, smoking marijuana or hash, and sniffing glue or gas. *Illicit substance use* was measured by summing the ratings of four binary items asking about lifetime use of

methamphetamine, ecstasy, drugs or medications prescribed to someone else, and other drugs not prescribed by a doctor including Vicodin, Percocet, or Oxycontin. Coefficient alphas were .79 and .76 for gateway illicit substance use, respectively.

Negative Substance Use Consequences Negative substance use consequences were measured with the *Young Adult Alcohol Problems Screening Test* (YAAPST; Hurlbut and Sher 1992). The YAAPST includes 12 items which assess the frequency with which individuals have experienced various negative consequences due to substance use (e.g., *I have said or done embarrassing things*). We modified the original scale to assess consequences the frequency (0 = never to 4 = four or more times) which have occurred in the past month due to alcohol, nicotine, marijuana, cocaine, inhalants, or other types of drugs ($\alpha = .91$).

Inhibitory Control and Attention A modified, computerized version of the *Stroop Task* (Siegrist 1995; MacLeod 1991) was created in E-Prime 3.0 (Psychology Software Tools (2016), Pittsburgh, PA) and presented on Acer netbooks. For the Stroop task, a series of color words (“red,” “blue,” or “yellow”) were sequentially presented for 2 s in the middle of a black screen. Participants were instructed to respond via button press to each word with the font color that it appeared in, not the name of the word itself. Colored stickers (yellow, blue, red) were placed on keyboard response keys (x, c, and v) to facilitate responding by reducing working memory load. Two trial types were presented—congruent and incongruent. On congruent trials, the word and color that it appeared in were the same (i.e., “red” appearing in red font). For incongruent trials, the word and font color were not the same (“red” appearing in blue font). Incongruent trials required participants to overcome their prepotent inclination to read the name of the word (thus engaging inhibitory control) and instead focus on and respond with the font color (attention). Sixty trials (thirty congruent and thirty incongruent) were randomly mixed and presented during the session. Ten practice trials were presented (prior to the appearance of color words) during which participants were asked to respond via button press with the correct color of a non-word prompt (“XXX”) to familiarize participants with response mode and mitigate practice effects. Outcome variables included reaction times (milliseconds) for congruent and incongruent trials and number of errors for congruent and incongruent trials.

Risk Taking A modified version of the *Balloon Analogue Risk Task* (BART; Lejuez et al. 2002) coded in E-Prime (Psychology Software Tools, Pittsburgh, PA) was also

administered to assess risk-taking propensity. During this task, participants were asked to inflate a virtual balloon (via keyboard button press) for potential monetary reward. Participants had the opportunity after each pump to stop and “bank” the earnings and move on to the next trial or continue to inflate the balloon. With each successive pump, the potential amount earned increased, however, so did the risk of the balloon popping. We used parameters for the BART as described in White et al. (2008): 30 balloons per run with an average break point of 64 out of a possible 128 pumps. For each pump, participants earned a “virtual” \$10. No monetary reward was actually delivered to the students. To incentivize performance, however, students earning the highest 3 total dollar amounts in a given classroom earned a tangible reward (i.e., assortment of university sports bags, travel mugs, candy). Standard outcome variables for the BART were collected, including the adjusted mean number of pumps (the average number of pumps on trials when the balloon did not explode), mean total monetary reward accrued, and mean number of balloon explosions (White et al. 2008).

Working Memory, Attention, and Emotion Regulation A modified *Emotional Faces N-back Task* (EFN-back), written in E-Prime (Psychology Software Tools, Pittsburgh, PA), was also administered as a means to assess working memory, attention, and emotion regulation (Ladouceur et al. 2005). During EFN-back administration, participants viewed a series of letters presented sequentially in the middle of the screen. Students were asked to respond via button press (spacebar) when the letter matched either a designated target letter (0-back condition) or matched the letter displayed 2 letters previous (2-back condition). On designated blocks, participants were presented with the series of letters flanked by pictures of emotional faces. Participants viewed eight blocks of trials flanked by emotional (happy, angry, or neutral) faces, as well as a block of trials with no faces present. Each block contained 12 trials, for a total of 96 trials (88 trials with faces). Outcome variables included the proportion of hits (correctly pressing the space bar when target letter appears) and false alarms (erroneously pressing the space bar in response to a non-target letter) across memory load (0-back, 2-back) and emotional valence conditions.

Engagement in Practice At post-test, students in the L2B condition were asked how often they practiced each of the seven program components since the beginning of the L2B program: body scan, three mindful breaths, mindful eating, mindfulness of thoughts, mindfulness of emotions, mindful movement and/or mindful walking, and loving-kindness practice. Each item was rated on an 8-point response scale from never to multiple times per day. The summary score was calculated by averaging over the 7 items ($\alpha = .96$).

Table 3 Student self-report scales by intervention status

	Pre		Post		Estimate	Intervention Effects		<i>d</i>
	Mean (SD)	Control	Mean (SD)	Control		<i>p</i>	<i>Adj-p</i>	
Internalizing								
Depression	5.91 (5.50)	5.27 (4.18)	5.57 (5.36)	4.89 (4.25)	0.45	0.36	0.99	0.09
Anxiety	5.39 (6.03)	4.27 (4.62)	4.99 (5.29)	4.6 (4.74)	−0.01	0.99	0.99	0.00
Rumination	3.34 (0.74)	3.27 (0.77)	3.29 (0.80)	3.11 (0.70)	0.17*	0.02	0.11	0.23
Stress of school performance	18.99 (6.65)	17.56 (6.13)	18.81 (6.72)	18.17 (6.49)	0.01	0.99	0.99	0.01
Stress of peer pressure	9.26 (5.62)	9.15 (4.58)	10.3 (5.66)	10.06 (5.37)	0.28	0.66	0.99	0.05
Somatization	27.9 (12.74)	26.77 (10.35)	28.32 (12.67)	28.01 (10.53)	0.21	0.86	0.99	0.02
Sleep								
Sleep quality	4.45 (0.97)	4.52 (0.83)	4.47 (1.03)	4.47 (0.77)	0.03	0.77	0.77	0.03
Falling asleep	4.01 (1.37)	4.18 (1.15)	4.04 (1.34)	4.14 (1.21)	−0.08	0.57	0.99	−0.06
Maintaining sleep	4.45 (1.01)	4.47 (0.97)	4.46 (1.12)	4.46 (0.80)	0.05	0.66	0.87	0.05
Reinitiating sleep	4.89 (1.09)	4.96 (0.92)	4.9 (1.07)	4.84 (0.94)	0.07	0.57	0.99	0.07
Emotion regulation								
Difficulties emot. reg.	2.76 (0.64)	2.66 (0.59)	2.75 (0.62)	2.74 (0.61)	−0.03	0.65	0.97	−0.04
Lack of emot. aware	2.99 (0.99)	2.98 (1.00)	2.82 (0.97)	3.09 (1.02)	−0.28*	0.01	0.03	−0.28
Lack of emotional clarity	2.34 (0.96)	2.42 (0.76)	2.37 (0.92)	2.52 (0.78)	−0.04	0.66	0.79	−0.06
Impulse control difficulties	2.95 (0.93)	2.67 (0.85)	2.86 (0.85)	2.75 (0.85)	−0.03	0.78	0.78	−0.02
Limited access to ER strat.	2.72 (0.80)	2.58 (0.74)	2.8 (0.72)	2.63 (0.66)	0.10	0.18	0.35	0.14
Goal-directed behavior diff.	2.92 (1.22)	2.71 (1.13)	3.16 (1.15)	2.69 (1.07)	0.43**	0.00	0.01	0.37
Mindfulness								
Mindfulness	25.41 (9.40)	25.24 (8.22)	24.59 (7.70)	24.48 (8.84)	−0.43	0.66	0.99	−0.05
Total self-compassion	36.93 (7.26)	38.18 (7.27)	36.88 (6.70)	37.8 (5.78)	−0.30	0.62	0.99	−0.04
Self-kindness	6.23 (1.77)	6.38 (1.63)	6.01 (1.77)	6.11 (1.71)	−0.05	0.79	0.99	−0.03
Self-judgment	5.54 (2.06)	5.43 (2.06)	5.48 (2.05)	5.45 (1.67)	0.03	0.91	0.91	0.01
Common humanity	5.67 (1.81)	5.83 (1.86)	5.82 (1.64)	5.81 (1.68)	0.08	0.68	0.99	0.04
Isolation	5.8 (2.14)	5.64 (2.14)	5.92 (2.14)	5.58 (1.88)	0.33	0.17	0.99	0.15
Mindfulness	6.67 (1.94)	6.95 (1.87)	6.44 (1.89)	6.55 (1.63)	0.04	0.66	0.99	0.02
Over-identified	6.28 (1.98)	5.9 (1.85)	5.99 (2.07)	5.57 (1.78)	0.24	0.28	0.99	0.13
Social connectedness	4.42 (1.16)	4.26 (0.99)	4.28 (1.15)	4.16 (1.01)	−0.03	0.80	0.99	−0.02
Mind-wandering	3.05 (1.54)	2.84 (1.24)	2.85 (1.46)	2.81 (1.48)	0.02	0.90	0.99	0.02
Growth mindset	4.07 (1.15)	4.06 (1.00)	4.14 (1.22)	4.01 (0.95)	0.12	0.33	0.99	0.11
Substance use								
Subst. use consequences	4.77 (7.04)	4.14 (5.7)	3.37 (6.23)	3.61 (5.85)				
Count portion of model					−0.05	0.78		0.95
Logistic portion of model					0.18	0.63		0.83
Substance use initiation	1.44 (1.69)	1.39 (1.73)	1.25 (1.65)	1.23 (1.72)	0.10	0.41		1.11
Illicit substance use	0.29 (0.78)	0.22 (0.65)	0.19 (0.64)	0.25 (0.75)	−0.56	0.19		0.57

* $p < .05$ ** $p < .01$, *Adj-p*, *p* value corrected using Benjamini-Hochberg procedure

Data Analyses

Main effects were analyzed using linear regression models for normally distributed continuous outcomes and generalized linear models for non-normally distributed count outcomes.

All outcome analyses were performed in *Mplus 7.2* (Muthén and Muthén 1998-2015). For self-report and neurocognitive/behavioral models and subsequent tests of moderation, only pre-test scores and a dummy variable representing school affiliation were included as covariates.

Effect size, d , was calculated by dividing the adjusted group mean difference by the pooled standard deviation. To correct for type I error due to multiple pairwise contrasts, p values were adjusted using a Benjamini-Hochberg correction (Benjamini and Hochberg 1995). As recommended, adjustments were made according to hypothesis families, with corrections made to measures grouped as follows: internalizing symptoms, measures of ER, measures of mindfulness and self-compassion, sleep, and EF sequentially (see Table 3).

Attrition and Missing Data Total attrition levels were low ($n = 3$; 1.6%). Examination of intervention by attrition interactions yielded no statistically significant differences on any pre-test variables. All missing data were handled using full-information maximum likelihood estimation (FIML) for each analysis under the assumption that missing is at random (Little and Rubin 2002).

Baseline Equivalence Subsequent to assignment, we assessed the baseline equivalence of intervention and

control groups and found no significant group differences in demographic characteristics. Only one baseline group difference was found for student self-report; impulse control difficulties were slightly higher in the intervention group, $t = -2.41$, $p < .05$. There were few differences in baseline neurocognitive outcomes; mean earnings on the BART task were higher in the control group ($t = 2.44$, $p < .05$); in the EFN-back task, the proportion of false alarm was higher in the intervention group for combination of 0-back and 2-back ($t = -3.02$, $p < .01$), for 0-back only ($t = -2.19$, $p < .05$), and for 2-back only ($t = -2.52$, $p < .05$; see Table 4).

Examining Clustering As students are nested in classrooms, we examined the intraclass correlation coefficient (ICC) values of each outcome and found that ICC values were in the trivial range (.01 to .05) (Muthen and Satorra 1995). Given the limited number of schools ($n = 2$), we include a dummy indicator of schools as a fixed effect to hold constant all unobserved characteristics that may vary between schools.

Table 4 Neurocognitive (EF) scales by intervention status

	Pre Mean (SD)		Post Mean (SD)		Estimate	Intervention Effects		d
	Intervention	Control	Intervention	Control		p	Adj- p	
STROOP								
RT of correct								
Congruent	562.46 (157.8)	544.36 (125.6)	537.97 (120.96)	559.44 (142.33)	-34.35*	0.01	0.04	-0.24
Incongruent	633.73 (171.47)	606.28 (137.39)	597.36 (134.39)	611.12 (157.29)	-29.81*	0.03	0.06	-0.19
Count of errors								
Congruent	1.28 (1.56)	1.25 (1.68)	1.28 (1.36)	1.45 (1.83)	-0.20	0.29	0.39	-0.12
Incongruent	4.22 (3.35)	3.68 (3.83)	3.71 (3.46)	3.71 (3.23)	-0.23	0.57	0.57	-0.06
BART								
Total points (Std)	0.15 (0.99)	0.19 (0.99)	0.04 (0.96)	0.04 (1.05)	0.04	ns	ns	0.04
# of explosions	7.48 (3.68)	8.21 (3.34)	10.08 (4)	10.76 (4.43)	-0.44	ns	ns	-0.12
Mean # of pumps	29.78 (12.63)	31.93 (13.37)	39.74 (14.52)	42.57 (15.53)	-1.77	ns	ns	-0.14
EFNBACK								
Proportion of hits								
All trials	0.66 (0.18)	0.66 (0.16)	0.69 (0.22)	0.71 (0.21)	0.00	ns	ns	-0.14
0-back only	0.85 (0.17)	0.87 (0.15)	0.8 (0.21)	0.82 (0.2)	-0.01	0.65	0.82	-0.07
2-back only	0.47 (0.27)	0.46 (0.23)	0.57 (0.28)	0.59 (0.26)	-0.04	0.93	0.93	-0.14
Proportion of false alarms								
All trials	0.05 (0.07)	0.03 (0.03)	0.05 (0.07)	0.03 (0.03)	-0.02**	0.00	0.02	-0.35
0-back only	0.05 (0.06)	0.04 (0.04)	0.04 (0.04)	0.05 (0.07)	-0.01*	0.01	0.03	-0.24
2-back only	0.05 (0.1)	0.02 (0.04)	0.03 (0.06)	0.05 (0.08)	-0.02†	0.09	0.15	-0.27

† $p < .10$

* $p < .05$

** $p < .01$, ns not significant, Adj- p p value corrected using Benjamini-Hochberg procedure

Results

Preliminary Analyses

Descriptive and Distributional Characteristics Prior to analyses, variables were examined to confirm approximate normality, identify outliers, homogeneity of variance, or any unusual patterns of missing data at the group or item level using SPSS 25.0. No unusual missing item patterns or distributional problems were detected.

As expected, self-report measures that include dichotomous or frequency counts of substance use were strongly skewed with many zero counts. We assessed over-dispersion by testing a dispersion parameter (α) and comparing goodness of fit between Poisson and negative binomial (NB) models (Atkins and Gallop 2007; Cameron and Trivedi 2013). If the Poisson assumption of equidispersion was violated, we estimated the NB model. We also assessed zero-inflation negative binomial models (ZINB) by examining frequency statistics and comparing model fits between the standard Poisson or NB and their zero-inflated counterparts. ZINB models assume two different origins of zero counts: the one due to a structural reason, and the other due to sampling. As a result, ZINB models consist of two parts, the one predicting a zero vs. a nonzero class (i.e., logistic portion) and the other modeling predicted counts for a nonzero class (i.e. counts portion). We used likelihood ratio (LR) tests to compare a set of nested models, e.g., Poisson vs. NB models, and relied on the Bayesian information criterion (BIC) to compare a set of non-nested models, e.g., Poisson vs. zero-inflated models (Green 1994; Hilbe 2011). Based on these analyses, we determined the most appropriate model for each of the count outcomes.

For gateway substance use, the standard Poisson provided the best fit; the dispersion parameter was not different from zero ($\alpha = 0, p > .05$). For illicit substance use, the standard NB model was the best model with the LR test favoring the NB over the Poisson ($\chi^2(1) = 23.99, p < .001$). For substance use consequences, we found evidence of significant overdispersion ($\alpha = 2.34, p < .001$) and the LR test favored the NB vs. the Poisson model, $\chi^2(1) = 552.28, p < .001$. The comparison of BIC fit statistics indicated stronger statistical support for the ZINB (BIC = 2164) over the standard NB (BIC = 2179). It was also substantively adequate to assume that zero observations in this measure have structural (i.e., those who reported zero because they have not been involved in substance use) and sampling origins (i.e., those who might be substance users but had no consequences).

L2B Impact on Student Self-Report Measures Table 3 presents the means, standard deviations, and main effect results for self-report scales. On nearly all the self-report measures, we found no significant effect of the intervention. Only 3 comparisons

yielded significant differences. At post-test, intervention students reported significantly lower levels of lack of emotional awareness ($t = -2.58, p = 0.01, d = -0.28$), an effect which remained significant after adjusting for multiple pairwise contrasts ($Adj-p = 0.03$). Contrary to our expectations, however, students in the intervention group reported higher levels of rumination ($t = 2.35, p = 0.02, d = 0.23$) and higher levels of difficulties engaging in goal-directed behavior ($t = 3.26, p = 0.01, d = 0.37$) at post-test. Although group differences in rumination were no longer significant after adjusting for multiple pairwise contrasts ($Adj-p = 0.11$), reported group differences related to difficulties engaging in goal-directed behavior remained statistically significant ($Adj-p = 0.01$).

Analyses of count models revealed no significant main effects of the intervention on substance use outcomes. Although it did not reach statistical significance in pre-post group comparisons, illicit substance use was considerably lower among students in the intervention group than those in the control ($t = -1.30, p = 0.19$). The estimated coefficient corresponds to the odds ratio of 0.57, which indicates that the intervention decreased illicit substance use by 43%.

L2B Impact on Behavioral Measures of EF Table 4 presents the means, standard deviations, and main effect results for the behavioral assessments of EF. Several statistically significant effects emerged favoring the intervention group. Although no statistically significant effects were found on BART measures, the intervention group outperformed the comparison group on several dimensions of both the Stroop and N-back tasks.

On the Stroop task, reaction times for correctly executed congruent and incongruent trials were lower for intervention students compared to controls with corresponding effect sizes of $-0.24 (t = -2.58, p = 0.01)$ and $-0.19 (t = -2.16, p = 0.03)$, respectively. After adjusting for multiple pairwise contrasts, the reaction time on congruent trails remained significant ($Adj-p = 0.04$) and incongruent trails were marginally significant ($Adj-p = 0.06$). There were no significant intervention effects related to the number of errors on the Stroop task.

On the emotional faces N-Back test, no significant effects were observed with respect to face valence (happy, angry, neutral) across groups. Consequently, we collapsed across valence and assessed the effect of memory load (0-back vs. 2-back) across group. Intervention students showed a significantly lower proportion of false alarms, with corresponding effect sizes of -0.35 for all trials ($t = -2.92, Adj-p = 0.02$), 0.24 for 0-back trials ($t = -1.69, Adj-p = 0.03$), and $-0.27 (t = -2.47, Adj-p = 0.15)$ for 2-back trials. There were no condition effects on proportion of hits.

Moderating Effects of Program Practice We conducted exploratory tests of moderation to determine whether different levels of out-of-class program practice (i.e., dosage) impacted

outcomes. The literature is limited with regards dose-response relationships in adolescent MBPs. Therefore, based on theory and in consultation with the program developer, we categorized practice into two discrete levels. Students with practice scale scores lower than 2 were categorized as an inadequate practice group (66% of intervention students), which corresponded to practicing less than once a month. Students with practice scale scores of 2 or higher were categorized as an adequate practice group (34% of intervention students), which corresponded to practicing at least once a month. The moderation analyses added the interaction term between group status

and program practice and removed the main effect of program practice, thereby allowing the levels of the outcomes to differ by program practice only among the intervention group. We chose an alpha level of $p < .05$ for main effects and of $p < .10$ for interaction effects due to the reduced power to detect interaction effects as suggested by Aguinis (1995).

Table 5 presents the moderation results of program practice on self-report outcomes. More frequent program practice tended to be associated with better outcomes for a variety of measures. As displayed in Fig. 3, students in the adequate practice group showed significantly lower levels of overall difficulties in ER ($d = 0.36$), lack of emotional awareness ($d = 0.35$), lack of emotional clarity ($d = 0.32$), impulse control difficulties ($d = 0.29$), mind-wandering ($d = 0.31$), and significantly higher levels of social connectedness ($d = 0.28$). An odds ratio (OR) of 0.75 indicating a 25% reduction in the odds of students engaging in gateway substance use. Although falling just short of reaching the .10 level of significance, there was a similar trend for practice effects for stress of school performance ($d = -0.27$), self-compassion ($d = 0.19$), and substance use consequences. To assess whether practice effects were due to existing pre-test difference, we compared adequate and inadequate practice groups on all pre-test self-report measures and found fewer differences than would be expected by chance. Thus, practice effects do not appear to be related to pre-test differences in “risk.”

The analysis of the ZINB model showed a trend for practice effects on substance use consequences for both count and logit portions of the model. The expected number of substance use consequences, among those likely in the non-zero group, was 34% lower for the adequate vs. inadequate practice groups (OR = 0.66). The probability of having no substance use consequences was 2.5 times higher for the adequate vs. inadequate practice groups (OR = 2.51). There were no significant interaction effects with practice for neurocognitive outcomes, indicating that the intervention impacts on neurocognitive outcomes were not significantly different by program practice levels.

Discussion

Our goal was to assess the efficacy of L2B as implemented by typical high school health teachers on measures of youth self-reports of emotion regulation, stress, anxiety and depression, substance use, and indicators of well-being. In addition, we conducted performance-based assessments of neurocognitive abilities using standardized EF measures. Prior studies have found youth exposed to the L2B program report significant improvements in emotion regulation, and concurrent decreases in negative affect, perceived stress, and stress-related somatic symptoms (Metz et al. 2013). Our study sought to extend these findings by conducting an independent study of

Table 5 Tests of moderation at adequate vs. inadequate practice levels

	Estimate	<i>d</i>
Depression	0.54	0.11
Anxiety	-0.23	-0.04
Rumination	0.14	0.19
Stress of school performance	-1.75	-0.27
Stress of peer pressure	-0.76	-0.15
Somatization	-2.16	-0.19
Sleep quality	0.13	0.14
Falling asleep	0.13	0.10
Maintaining sleep	0.05	0.05
Reinitiating sleep	0.21	0.21
Difficulties-Emotion regulation	-0.22*	-0.36
Lack of emotional awareness	-0.35*	-0.35
Lack of emotional clarity	-0.28*	-0.32
Impulse control difficulties	-0.26†	-0.29
Limited access to emotion regulation strategies	-0.11	-0.15
Difficulties engaging in goal-directed behavior	0.17	0.15
Mindfulness	1.87	0.21
Total self-compassion	1.41	0.19
Self-kindness	0.28	0.17
Self-judgment	-0.28	-0.14
Common humanity	0.04	0.02
Isolation	0.07	0.03
Mindfulness	0.43	0.23
Over-identified	-0.13	-0.07
Social connectedness	0.31†	0.28
Mind-wandering	-0.44†	-0.31
Growth mindset	0.188	-0.12
Substance use consequences		
Count portion of model	-0.42	0.66
Logistic portion of model	0.92	2.51
Gateway substance use	-0.28	0.75
Illicit substance use	0.76	2.14

† $p < .10$

* $p < .05$

** $p < .01$

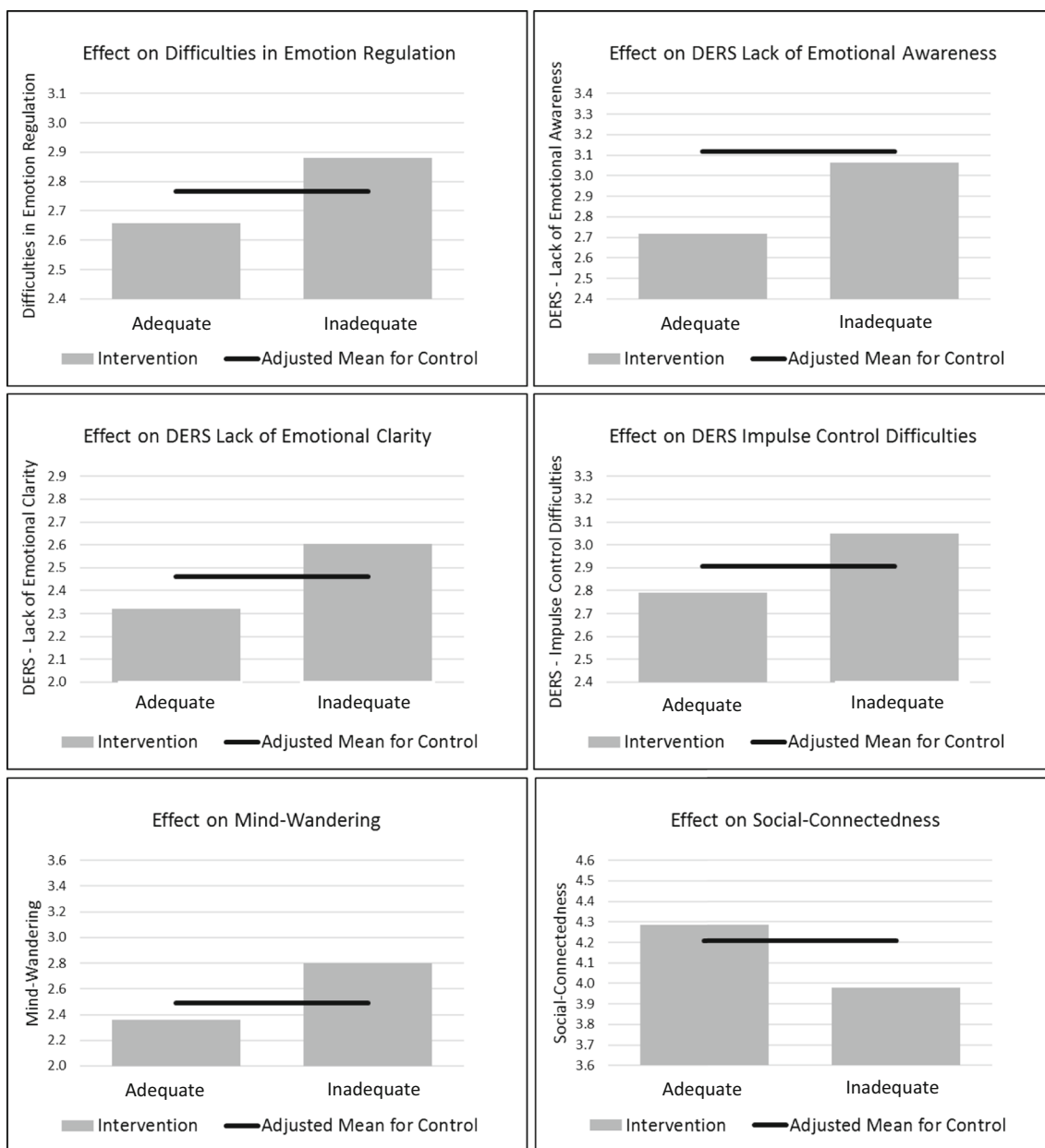


Fig. 3 Differential intervention impacts by program practice levels, * $p < .05$, ** $p < .01$

L2B delivered by teachers in a real-world, ethnically diverse school setting.

Contrary to our expectations, we did not find intervention main effects on most of the adolescent self-report measures. Thus, we were not able to replicate findings from previous studies of L2B on these measures of depression (negative affect), emotion regulation, or perceived stress shown in some previous smaller studies of L2B. However, we did find that students exposed to L2B showed significant improvements on some components of 2 of the 3 measures of EF. Intervention students showed a significantly higher level of selective attention and inhibitory control as reflected by performance on the Stroop and EFN-back tasks, respectively. Across most

variants of the Stroop task, slower and/or less accurate responding during the incongruent vs. congruent condition reflects a greater internal processing demand to resolve stimulus/response conflict and inhibit prepotent responding (Wolf et al. 2014). Both intervention and control groups showed this robust effect of greater demand on incongruent trials. However, students in the intervention, but not control group, generated significantly faster correct responses during the post-test on both congruent and incongruent trials, controlling for pre-test latencies. These data suggest that students who participated in L2B showed an overall higher level of selective attention to the task regardless of trial type. It may be the case that an increased focus on the task at hand coupled

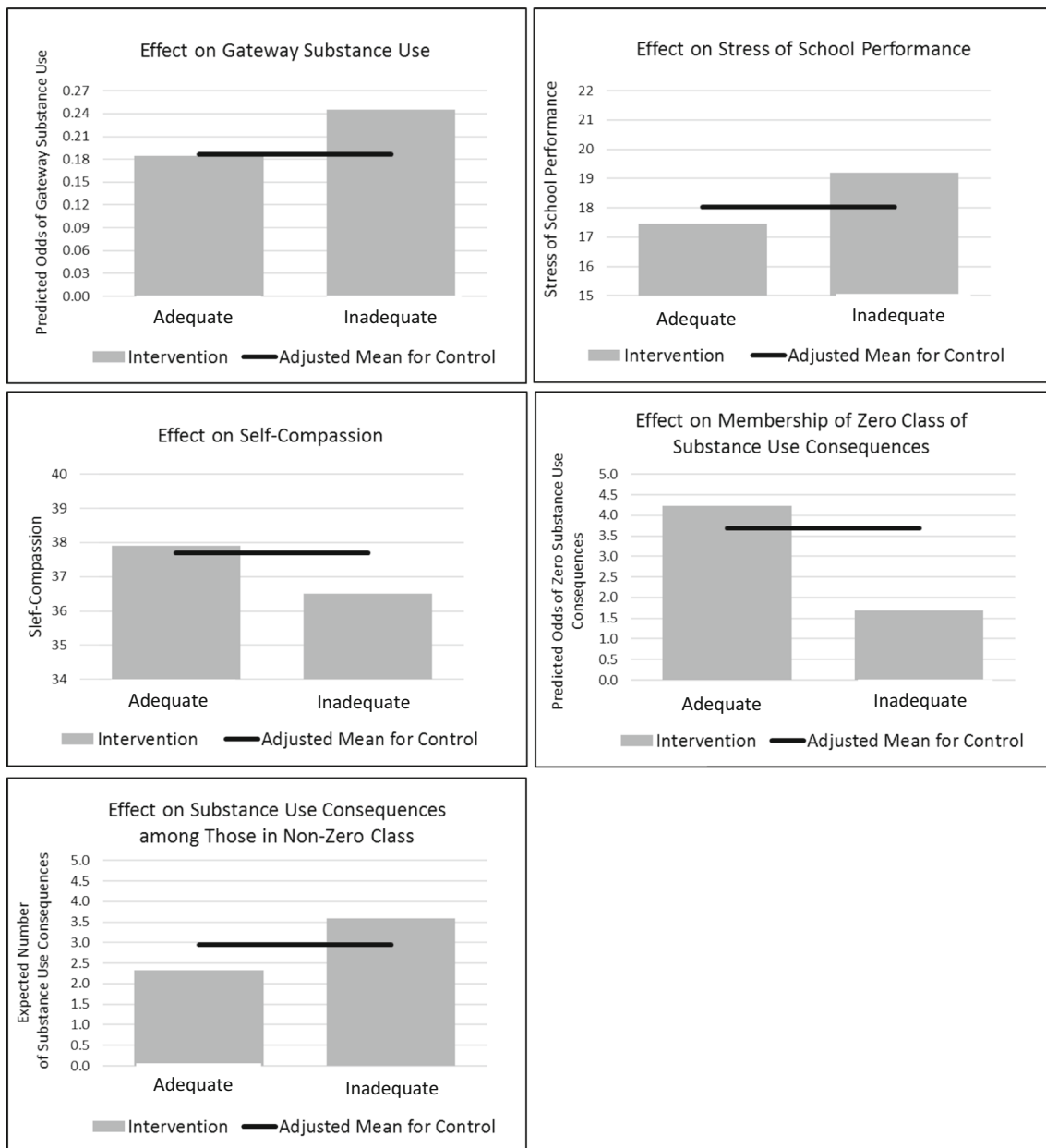


Fig. 3 continued.

with decreased attention to distracters (e.g., noise/movement in the room), promoted by intervention training, lessened the overall cognitive load on students’ interference, and inhibitory control related brain circuitry, thus improving task performance.

On the emotional faces N-back task, we found no effect of face valence on task performance. When collapsed across valence type to assess the effect of memory load, however, we found that the intervention group generated a lower proportion of false alarms on 0-back trials and a trend (after adjusting for multiple tests) on the 2-back trials. A false alarm on the 2-back task occurs when a participant incorrectly presses the response button (i.e., current letter does not match the letter shown two

screens back); it is sometimes referred to as an error of commission. This measure can be interpreted as failure to pay attention to or encode/update the internal representation of the target letter. Intervention participants thus appear to be paying more attention to changes in stimuli characteristics (i.e., the varying letters). Notably, this occurs in the absence of significant changes in correct response rate (i.e., “hits”) or, by extension, the number of misses (inverse of hits). Previous literature suggests that errors of omission and commission represent distinct error types supported by different psychological processes and likely different underlying neural circuitry (Meule 2017). Thus, it may be the case that the intervention was particularly effective for a specific dimension of

attentional control related to commission errors. It is also particularly notable that this difference was observed during 2-back trials, as previous work suggests that attention can be disrupted under conditions of higher working memory load (e.g., Judah et al. 2013). Considered together, across the Stroop and EFN-back tasks, the L2B intervention led to small but consistent improvements in tests of selective attention.

In addition, the results of the moderation analyses suggest that beneficial program effects within the treatment group are influenced by the degree to which students report actively engaging in mindfulness practices outside of the classroom setting. Specifically, we found that students who reported higher rates of practice (greater than once per month) showed significant small-to-moderate interaction effects on measures of emotional awareness, emotional clarity, impulse control difficulties, social connectedness, mind-wandering, substance use, stress reduction, and self-compassion. A similar but non-significant trend was found for stress of school performance and self-compassion. On substance use measures, those in the adequate as compared to the inadequate practice group had an odds ratio that indicated there were 2.5 times more likely to report having no substance use-related consequences. Unlike self-report measures, we found no significant interaction effects with practice for EF outcomes suggesting that direct effects were not conditional on dosage in the same manner as self-reported social-emotional and behavioral outcomes. Although no baseline characteristics predicted practice time, we note in that practice time was not experimentally varied and thus should be subject to replication.

In considering the mixed results of this trial in light of prior findings, several design considerations are worth noting. First, this is one of the first studies of mindfulness-based interventions in education in which the intervention was presented by existing teachers rather than highly trained mindfulness experts. Our attempt here was to train and utilize regular classroom teachers who were relatively naïve regarding mindfulness at the onset of the study. The question was whether or not we could provide regular classroom teachers with both the background training in personal mindfulness as well as training in the implementation of L2B for students necessary to impact social-emotional and behavioral outcomes. Given the fact that L2B is a relatively well-structured and manualized curriculum, it represents an ideal program in which to explore these research questions.

We provided extensive supervision with both occasional live observations as well as weekly feedback calls after reviewing videos of previous lessons. However, none of the intervention teachers achieved highly proficient levels of implementation fidelity despite these supports. Although fidelity of implementation was not perfect, it likely provides an “average case” test of probable outcomes when utilizing teachers currently working in a typical American high school setting which has significant

implications for scaling potential. Our experience suggests that it is not easy to train high school teachers to deliver this kind of mindfulness intervention without several years of support, and that this represents a key on-going question of implementation for such programs. In addition, detecting change in clinical symptoms in universal populations can be challenging given the vast majority of students participating in such interventions are unlikely to be symptomatic (Greenberg and Abenavoli 2017).

Second, there is little literature to guide program developers regarding the needed dosage to have significant impact on typical adolescents. Studies are needed that vary both dosage (number of sessions) and density of sessions (number of times per week) to assess how to optimize outcomes. Third, as reported above, most students did not report high levels of practice outside the classroom in spite of supplying audio-guided practices that could be used on a regular basis. We concur with Bailey et al.’ (2018) commentary regarding the practical challenges of implementing mindfulness-based practices at scale in school settings and that many students have limited “capacity to take on optional extras” such as mindfulness.

Although the present study did integrate mindfulness practice into the regular school day as recommended by Bailey et al. (2018), our findings suggest that adolescents’ own choices and willingness to engage in practices outside of the regular school setting may be key to the success of universal school-based approaches. As amount of practice outside of the structured school setting showed significant effects on outcomes, creating higher motivation to use practices on a regular basis should be considered a high priority for program development. Likewise, students may need direction and support on how to fit mindfulness-based practices into their busy schedules in order to form a regular practice habit.

Limitations and Directions for Future Research

Although this study had several strengths, including objective behavioral measures, authentic intervention implementation by trained classroom teachers, and a diverse sample of students, it is not without limitations. Not all subscale measures demonstrated optimal reliability in this field trial, and therefore, findings related to these subscales should be interpreted with caution. Not all measures were specifically designed for use with adolescents, which presents another potential limitation. Although objective behavioral data was collected (e.g., EF), the study featured youth self-report of social-emotional functioning and behavior as opposed to parent or teacher report which may yield alternative findings. Although there was a relatively small amount of missing data and it was not related to condition our missing data analysis assumes, it was missing at random. Moreover, finding new ways to measure practice that do not rely upon participant self-reports, including direct measures of usage through i-App delivered home practices,

may be one way to improve our understanding of dose-response relations in these kinds of program for adolescents. In addition, given the intervention was delivered in a universal classroom setting, we were unable to control for the number of sessions students were exposed to due to student absences. We relied on a relatively small, but well-trained and supervised sample of teachers to examine the potential effectiveness of the L2B program on adolescent outcomes. It is quite possible that specific teacher or classroom qualities may differentially impact the effectiveness of MBPs delivered in this manner. Finally, the present study focused exclusively on pre- and post-intervention effects, and did not include a longer-term longitudinal investigation of effects.

This study contributes to the growing body of research on the effects of mindfulness programs delivered in school settings. While some studies have shown significant effects (Kuyken et al. 2013; Metz et al. 2013; Raes et al. 2014; Sibinga et al. 2016), others have not (Johnson et al. 2016, 2017). Given the mixed nature of findings in this study, we propose that a logical next step in the evaluation of such programs is to conduct larger effectiveness trials with a closer examination of dosage and practice effects (Kuyken et al. 2017). Examination of contextual, attitudinal, and motivational factors predicting adolescent adoption of mindfulness intervention strategies outside of school settings, and qualitative and mixed methods research investigating how best to support this, is a particularly important goal for the next wave of school-based mindfulness interventions.

Despite these limitations, this study provides mixed support regarding the potential effectiveness of a universal mindfulness program for high school students. The absence of direct effects on self-report measures implies that simply exposing adolescents to a mindfulness curriculum within the context of typical instruction, in the absence of supports for implementation, is unlikely to substantially impact youth self-report of social-emotional well-being or behavior. However, changes on EF favoring the intervention group were noted suggesting possible benefits on tasks related to susceptibility to cognitive interference, and selective attention are possible. Tests of moderation revealed dosage effects such that students who adopt the mindfulness practices they are taught and use them somewhat regularly can indeed benefit on multiple fronts. Greater effects may be possible with higher levels of dosage, and utilization of practices outside of the immediate school setting. As such, future research examining optimal dosage and strategies to increase strategy utilization is a key priority for future research.

Authors' Contribution JF, PB, MG, and DS designed and executed the study, and wrote the paper. YO analyzed study data and summarized results. CG assisted with the design and analysis of executive function measures. JM assisted with recruitment, consent, and data collection. PB and EB assisted with intervention implementation, professional

development, and coaching activities. KK, JM, and JL assisted with data collection. RR collaborated in the writing and editing of the final manuscript.

Funding The project described was supported by Award Number R305A140113 from the Institute of Educational Sciences (IES). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Institute of Educational Sciences or the U. S. Department of Education.

Declarations

Ethics Approval and Consent to Participate All procedures performed in the study involving human participants were in accordance with the ethical standards of the Pennsylvania State University Institutional Review Board and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Conflict of Interest In accordance with ethical obligations, Patricia C. Broderick, developer of Learning to Breathe, is reporting a financial interest that may be affected by the research reported in the enclosed paper. The remaining authors declare that they have no conflict of interest.

References

- Aguinis, H. (1995). Statistical power with moderated multiple regression in management research. *Journal of Management*, *21*, 1141–1158.
- Atkins, D. C., & Gallop, R. J. (2007). Rethinking how family researchers model infrequent outcomes: a tutorial on count regression and zero-inflated models. *Journal of Family Psychology*, *21*, 726–735.
- Bailey, N. W., Chambers, R., Wootten, A., & Hassed, C. S. (2018). Commentary Regarding Johnson et al. (2017) "A randomized controlled evaluation of a secondary school mindfulness program for early adolescents: do we have the recipe right yet?". *Mindfulness*, *9*, 1668–1670.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*(1), 289–300.
- Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: changes and correlates. *Developmental Review*, *29*, 180–200.
- Black, D. S., & Slavich, G. M. (2016). Mindfulness meditation and the immune system: a systematic review of randomized controlled trials. *Annals of the New York Academy of Sciences*, *1373*(1), 13–24.
- Blakemore, S. J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, *65*, 187–207.
- Bluth, K., Campo, R. A., Pruteanu-Malinici, S., Reams, A., Mullarkey, M., & Broderick, P. C. (2016). A school-based mindfulness pilot study for ethnically diverse at-risk adolescents. *Mindfulness*, *7*, 90–104.
- Broderick, P. C. (2013). *Learning to BREATHE: a mindfulness curriculum for adolescents to cultivate emotion regulation, attention, and performance*. Oakland: New Harbinger Publications.
- Broderick, P. C., & Metz, S. (2009). Learning to BREATHE: a pilot trial of a mindfulness curriculum for adolescents. *Advances in School Mental Health Promotion*, *2*, 35–46.
- Caballero, A., Granberg, R., & Tseng, K. Y. (2016). Mechanisms contributing to prefrontal cortex maturation during adolescence. *Neuroscience and Biobehavioral Reviews*, *70*, 4–12.
- Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (Vol. 53). Cambridge: Cambridge University Press.

- Carsley, D., Khoury, B., & Heath, N. L. (2018). Effectiveness of mindfulness interventions for mental health in schools: a comprehensive meta-analysis. *Mindfulness*, *9*, 693–707.
- Casey, B. J. (2015). Beyond simple models of self-control to circuit-based accounts of adolescent behavior. *Annual Review of Psychology*, *66*, 295–319.
- Chambers, R., Gullone, E., & Allen, N. B. (2009). Mindful emotion regulation: An integrative review. *Clinical Psychology Review*, *29*, 560–572.
- Cohen-Gilbert, J. E., & Thomas, K. M. (2013). Inhibitory control during emotional distraction across adolescence and early adulthood. *Child Development*, *84*, 1954–1966.
- Crane, R. S., Brewer, J., Feldman, C., Kabat-Zinn, J., Santorelli, S., Williams, J. M. G., & Kuyken, W. (2016). What defines mindfulness-based programs? The warp and the weft. *Psychological Medicine*, *47*, 990–999.
- Dahl, R. E. (2004). Adolescent brain development: a period of vulnerabilities and opportunities. Keynote address. *Annals of the New York Academy of Sciences*, *1021*(1), 1–22.
- Davidson, R. J. (2000). Affective style, psychopathology, and resilience: brain mechanisms and plasticity. *The American Psychologist*, *55*, 1196–1214.
- Dreyfuss, M., Caudle, K., Drysdale, A. T., Johnston, N. E., Cohen, O., Somerville, L. H., Galvan, A., Tottenham, N., Hare, T. A., & Casey, B. J. (2014). Teens impulsively react rather than retreat from threat. *Developmental Neuroscience*, *36*, 220–227.
- Dweck, C. S. (1999). *Self-theories: their role in motivation, personality, and development*. Philadelphia: Psychology Press.
- Felver, J. C., Celis-de Hoyos, C. E., Tezanos, K., & Singh, N. N. (2016). A systematic review of mindfulness-based interventions for youth in school settings. *Mindfulness*, *7*, 34–45.
- Felver, J. C., Clawson, A., Morton, M., Brier-Kennedy, Janack, P., & DiFlorio, R. (2018). School-based mindfulness intervention supports adolescent resiliency. *International Journal of School and Educational Psychology*, *7*, 111–122.
- Fung, J., Kim, J. J., Jin, J., Chen, G., Baer, L., & Lau, A. (2019). A randomized trial evaluating school-based mindfulness intervention for ethnic minority youth: exploring mediators and moderators of intervention effects. *Journal of Abnormal Child Psychology*, *47*, 1–19.
- Gottfried, J. A. (Ed.). (2011). *Neurobiology of sensation and reward*. Boca Raton: CRC Press/Taylor & Francis.
- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, *26*, 41–54.
- Greco, L. A., Baer, R. A., & Smith, G. T. (2011). Assessing mindfulness in children and adolescents: development and validation of the child and adolescent mindfulness measure (CAMM). *Psychological Assessment*, *23*, 606–614.
- Green, W. (1994). *Accounting for excess zeros and sample selection in Poisson and negative binomial regression models*. (Working paper EC-94-10). New York: New York University, Department of Economics.
- Greenberg, M. T., & Abenavoli, R. (2017). Universal interventions: fully exploring their impacts and potential to produce population-level impacts. *Journal of Research on Educational Effectiveness*, *10*, 40–67.
- Greenberg, M. T., & Lippold, M. A. (2013). Promoting healthy outcomes among youth with multiple risks: innovative approaches. *Annual Review of Public Health*, *34*, 253–270.
- Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: conceptual foundations. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 3–24). New York: The Guilford Press.
- Hare, T. A., Tottenham, N., Galvan, A., Voss, H. U., Glover, G. H., & Casey, B. J. (2008). Biological substrates of emotional reactivity and regulation in adolescence during an emotional go-nogo task. *Biological Psychiatry*, *63*, 927–934.
- Hilbe, J. M. (2011). *Negative binomial regression* (2nd ed.). Cambridge University Press.
- Hurlbut, S. C., & Sher, K. J. (1992). Assessing alcohol problems in college students. *Journal of American College Health*, *41*(2), 49–58.
- Johnson, C., Burke, C., Brinkman, S., & Wade, T. (2016). Effectiveness of a school-based mindfulness program for transdiagnostic prevention in young adolescents. *Behaviour Research and Therapy*, *81*, 1–11.
- Johnson, C., Burke, C., Brinkman, S., & Wade, T. (2017). A randomized controlled evaluation of a secondary school mindfulness program for early adolescents: do we have the recipe right yet? *Behaviour Research and Therapy*, *99*, 37–46.
- Judah, M. R., Grant, D. M., Mills, A. C., & Lechner, W. V. (2013). The neural correlates of impaired attentional control in social anxiety: an ERP study of inhibition and shifting. *Emotion*, *13*, 1096–1106.
- Kabat-Zinn, J. (1990). *Full catastrophe living: using the wisdom of your body and mind to face stress, pain and illness*. New York: Delacorte.
- Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B., Berry, J. T., & Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, *114*, 163–173.
- Kuyken, W., Weare, K., Ukoumunne, O. C., Vicary, R., Motton, N., Burnett, R., Cullen, C., Hennesly, S., & Huppert, F. (2013). Effectiveness of the mindfulness in schools programme: non-randomised controlled feasibility study. *The British Journal of Psychiatry*, *203*, 126–131.
- Kuyken, W., Nuthall, E., Byford, S., Crane, C., Dalgleish, T., Ford, T., Greenberg, M. T., Ukoumunne, O. C., Viner, R. M., Williams, J. M. G., Brett, D., Lord, L., Palmer, L., Sonley, A., Taylor, L., Raja, A., & Casey, T. (2017). The effectiveness and cost-effectiveness of a mindfulness training programme in schools compared with normal school provision (MYRIAD): study protocol for a randomised controlled trial. *Trials*, *18*, 194.
- Ladouceur, C. D., Dahl, R. E., Williamson, D. E., Birmaher, B., Ryan, N. D., & Casey, B. J. (2005). Altered emotional processing in pediatric anxiety, depression, and comorbid anxiety-depression. *Journal of Abnormal Child Psychology*, *33*(2), 165–177.
- Larson, R. W., Moneta, G., Richards, M. H., & Wilson, S. (2002). Continuity, stability, and change in daily emotional experience across adolescence. *Child Development*, *73*(4), 1151–1165.
- LeBourgeois, M. K., Giannotti, F., Cortesi, F., Wolfson, A. R., & Harsh, J. (2005). The relationship between reported sleep quality and sleep hygiene in Italian and American adolescents. *Pediatrics*, *115*, 257–265.
- Lee, R. M., Draper, M., & Lee, S. (2001). Social connectedness, dysfunctional interpersonal behaviors, and psychological distress: Testing a mediator model. *Journal of Counseling Psychology*, *48*(3), 310–318.
- Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, *8*(2), 75–84.
- Little, R. J., & Rubin, D. B. (2002). *Bayes and multiple imputation, Statistical Analysis with Missing Data, Second Edition* (pp. 200–220). Hoboken: John Wiley & Sons.
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin*, *109*, 163–203.
- Maynard, B. R., Solis, M. R., Miller, V. L., & Brendel, K. E. (2017). Mindfulness-Based interventions for improving cognition, academic achievement, behavior, and socioemotional functioning of primary and secondary school students. *Campbell Systematic Reviews* 2017: 5. *Campbell Collaboration*.
- Metz, S. M., Frank, J. L., Reibel, D., Cantrell, T., Sanders, R., & Broderick, P. C. (2013). The effectiveness of the learning to

- BREATHE program on adolescent emotion regulation. *Research in Human Development*, 10, 252–272.
- Meule, A. (2017). Reporting and interpreting working memory performance in n-back tasks. *Frontiers in Psychology*, 8, 1–3.
- Mrazek, M. D., Phillips, D. T., Franklin, M. S., Broadway, J. M., & Schooler, J. W. (2013). Young and restless: validation of the Mind-Wandering Questionnaire (MWQ) reveals disruptive impact of mind-wandering for youth. *Frontiers in Psychology*, 4, 1–7.
- Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide* (Seventh ed.). Los Angeles, CA: Muthén & Muthén.
- Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, 25, 267–316.
- Nook, E. C., Sasse, S. F., Lambert, H. K., McLaughlin, K. A., & Somerville, L. H. (2018). The nonlinear development of emotion differentiation: granular emotional experience is low in adolescence. *Psychological Science*, 29, 1346–1357.
- Powers, A., & Casey, B. J. (2015). The adolescent brain and the emergence and peak of psychopathology. *Journal of Infant, Child, and Adolescent Psychotherapy*, 14, 3–15.
- Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from <http://www.pstnet.com>.
- Raes, F., Pommier, E., Neff, K. D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the self-compassion scale. *Clinical Psychology & Psychotherapy*, 18, 250–255.
- Raes, F., Griffith, J. W., Van der Gucht, K., & Williams, J. M. G. (2014). School-based prevention and reduction of depression in adolescents: a cluster-randomized controlled trial of a mindfulness group program. *Mindfulness*, 5, 477–486.
- Riediger, M., Schmiedek, F., Wagner, G. G., & Lindenberger, U. (2009). Seeking pleasure and seeking pain: differences in prohedonic and contra-hedonic motivation from adolescence to old age. *Psychological Science*, 20, 1529–1535.
- Roeser, R. (2016). Process of teaching, learning, and transfer in mindfulness-based interventions (MBIs) for teachers: a contemplative educational perspective. In K. A. Schonert-Riechl & R. W. Roeser (Eds.), *Handbook of Mindfulness in Education* (pp. 149–170). New York: Springer.
- Rothbart, M. K., Sheese, B. E., & Posner, M. I. (2007). Executive attention and effortful control: linking temperament, brain networks, and genes. *Child Development Perspectives*, 1(1), 2–7.
- Sibinga, E. M., Webb, L., Ghazarian, S. R., & Ellen, J. M. (2016). School-based mindfulness instruction: an RCT. *Pediatrics*, 137(1), 1–8.
- Siegrist, M. (1995). Reliability of the Stroop test with single-stimulus presentation. *Perceptual and Motor Skills*, 81(3_suppl), 1295–1298.
- Snyder, H. R., & Hankin, B. L. (2016). Spiraling out of control: stress generation and subsequent rumination mediate the link between poorer cognitive control and internalizing psychopathology. *Clinical Psychological Science: A Journal of the Association for Psychological Science*, 4(6), 1047–1064.
- Somerville, L. H., Jones, R. M., & Casey, B. J. (2010). A time of change: behavioral and neural correlates of adolescent sensitivity to appetitive and aversive environmental cues. *Brain and Cognition*, 72, 124–133.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, 166(10), 1092–1097.
- Spoth, R., Redmond, C., Shin, C., Greenberg, M., Clair, S., & Feinberg, M. (2007). Substance-use outcomes at 18 months past baseline: the PROSPER community–university partnership trial. *American Journal of Preventive Medicine*, 32(5), 395–402.
- Tang, Y. Y., Hölzel, B. K., & Posner, M. I. (2015). The neuroscience of mindfulness meditation. *Nature Reviews. Neuroscience*, 16, 213–225.
- Trapnell, P. D., & Campbell, J. D. (1999). Private self-consciousness and the five-factor model of personality: distinguishing rumination from reflection. *Journal of Personality and Social Psychology*, 76, 284–304.
- Walker, L. S., Beck, J. E., Garber, J., & Lambert, W. (2008). Children's Somatization Inventory: psychometric properties of the revised form (CSI-24). *Journal of Pediatric Psychology*, 34, 430–440.
- White, T. L., Lejuez, C. W., & de Wit, H. (2008). Test-retest characteristics of the Balloon Analogue Risk Task (BART). *Experimental and Clinical Psychopharmacology*, 16, 565–570.
- Wolf, D., Zschutschke, L., Scheurich, A., Schmitz, F., Lieb, K., Tusher, O., & Fellgiebel, A. (2014). Age-related increases in stroop interference: delineation of general slowing based on behavioral and white matter analysis. *Human Brain Mapping*, 35, 2448–2458.
- Zelazo, P. D. (2015). Executive function: Reflection, iterative reprocessing, complexity, and the developing brain. *Developmental Review*, 38, 55–68.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.