Yoga Improves Academic Performance in Urban High School Students Compared to Physical Education: A Randomized Controlled Trial

Marshall Hagins¹ and Andrew Rundle²

ABSTRACT—Yoga programs within schools have become more widespread but research regarding the potential effect on academic achievement remains limited. This study cluster-randomized 112 students within a single New York City public high school to participate in either school-based yoga or physical education (PE) for an entire academic year. The primary outcome was mean annual grade point average (GPA). Psychosocial variables (self-regulation, executive function, well-being, and mindfulness) were examined as mediators. The study’s primary hypothesis that yoga would improve academic performance was not supported by intent to treat analysis; however, a significant interaction was observed between class assignment and class participation. Among students with higher participation, those assigned to yoga classes had a significantly higher GPA. For example, at 49 classes of participation for both groups, students assigned to yoga classes had an estimated 2.70 higher mean GPA (effect size = 0.31) than students assigned to PE.

Success in school is a strong predictor for social and occupational status in adulthood (Ross & Wu, 1996; Schoenbaum & Waidmann, 1997) as well as for lifelong health and quality of life (Feinstein, 2002). School programs often focus on factors which have been shown to positively influence academic performance, including physical activity (Chomitz et al., 2009; Singh, Uijtdewilligen, Twisk, van Mechelen, & Chinapaw, 2012), health (Ickovic et al., 2014), social emotional learning (Bavarian et al., 2013), and self-regulation (SR) and executive function (EF) (Blair & Razza, 2007; Brock, Rimm-Kaufman, Nathanson, & Grimm, 2009; Denham, 2006; Li-Grining, Votruba-Drzal, Maldonado-Carreno, & Haas, 2010; Mashburn & Pianta, 2006; Valiente, Lemery-Chalfant, Swanson, & Reiser, 2008).

Self-regulation and EF in classroom settings are generally described as the management of behavioral, cognitive, emotional, and attentional resources relative to achieving a learning goal (Pintrich, 2000). More specifically, SR refers to a complex of acquired, intentional skills involved in controlling, directing, and planning one’s cognitions, emotions, and behavior (Schunk & Zimmerman, 1997) whereas the closely related construct EF refers to working memory, arousal, emotional control, problem solving, shifting activities appropriately, organizing, and self-monitoring (Gioia, Isquith, Guy, & Kenworthy, 2000).

Self-regulation and EF allow students to attend appropriately to classroom activities, remember information, inhibit distractors, and persist toward goals, all of which demonstrate successful behavioral regulation. Characteristics of SR and EF are positively related to reading, math, and linguistic abilities (Fabetes, Martin, Hanish, Anders, & Madden-Derdich, 2003; NICHD, 2003) and have been shown to improve the ability to remain focused on tasks (Zimmerman, 1998) and process detailed situations more accurately (Lemerise & Arsenio, 2000; NICHD, 2003).
and/or EF have been linked to positive academic outcomes (Graziano, Reavis, Keane, & Calkins, 2007; Howse, Calkins, Anastopoulos, Keane, & Shelton, 2003; Mazzocco & Kover, 2007; Miyake et al., 2000). Given the above, programs which promote SR and EF may provide benefit to academic performance.

Recently, school programs based on the practices of yoga have become more common (Little Flower Yoga, n.d.; Sonima Foundation, n.d.; Yoga Kids, n.d.). The explicit goal of many of these programs is to increase academic performance via changes in self-regulation and executive function. Yoga is a contemplative practice which combines physical activity, breathing exercises, meditation, and relaxation and is aimed at the control of mental processes related to self-regulation. The empirical evidence supporting yoga’s effects on self-regulation as well as potential neurobiological mechanisms have recently been described in detail (Gard, Noggle, Park, Vago, & Wilson, 2014). Several reviews suggest that yoga improves mental health and behavior in children and adolescents (Birdee et al., 2009; Galantino, Galbavy, & Quinn, 2008; Kaley-Isley, Peterson, Fischer, & Peterson, 2010). Recent reviews of the literature on yoga programs in schools suggest that yoga can enhance emotional balance, attentional control, cognitive efficiency and self-esteem, and decrease anxiety, negative thought patterns, and negative behavior (Ferreira-Vorkapic et al., 2015; Serwacki & Cook-Cottone, 2012).

Mechanisms linking yoga and potential benefits of increased academic performance are based on the demands of yoga practice, which requires volitional control and sustained focus. The goal of yoga is to practice regulation of behavior and awareness such that these skills become more automatic and accessible and can be applied within varied environments. The implicit hypothesis is that skills of SR and EF learned within yoga will transfer to the context of the classroom, enhancing academic performance. Yoga training may increase skills of SR and EF contributing to academic success; however, yoga does not explicitly train for improved academic performance. Consequently, yoga may positively influence what is known as “self-regulated learning” but does not specifically train aspects of that construct such as self-efficacy or motivation to attain specific educational goals.

To our knowledge, there are three quantitative studies that have examined the potential effects of yoga on the academic performance of otherwise healthy students using measures such as standardized tests or school grades. Two of the studies found yoga to have significant effects on academic performance compared to non active control (Kauts & Sharma, 2009) or an active control (Butzer, van Over, Noggle Taylor, & Khalsa, 2015), whereas the remaining study reported mixed effects (Smith, Connington, McQuillin, & Crowder Bierman, 2014). Given the recent increase in yoga programs in schools and the real-world relevance of academic performance as an outcome, there is a need for additional studies to clarify the potential for yoga to impact academic performance. Particularly needed are rigorous studies using active control groups, examination of the effects of yoga practice across the entire school year, measures of fidelity of the intervention, and exploration of potential mediating variables related to current theory about mechanisms of yoga. Therefore, the purpose of this study was to examine the effects of a year-long school-based yoga program on academic performance and explore potential mediating effects of emotional regulation and executive function.

MATERIALS AND METHODS

This study used cluster randomization at the level of class within a single school to compare the effects of a yoga program to a standard physical education (PE) program on academic performance. The hypothesis was that the yoga program would provide significantly improved academic performance (annual grade point averages [GPA]) compared to the PE class. Participation in class was expected to moderate the relationship, and levels of emotional regulation and executive function were expected to mediate the relationship. Secondary potential mediators were sense of well-being and mindfulness. A qualitative assessment of students and teachers providing information on challenges and strengths of the program was also performed but will be reported elsewhere. Prior to recruitment the study was approved by the Long Island University and New York City Department of Education Institutional Review Boards. The study was registered with Clinicaltrials.gov (NCT02329015).

Participants

The study sample was drawn from 9th, 10th, and 11th grade students in a single public high school in New York City (demographics: 11% Asian, 22% Black, 59% Hispanic, and 8% White). All students who were cleared for PE class were eligible for the study. During the student’s first advisory period at the start of the academic year (September, 2014–2015), the study was explained to the students and any questions were answered. At that time letters explaining the study, as well as consent for the child’s participation, were given to students to bring home for the parents to fill out. An assent form for the child to fill out was also included. The parent or the child was asked to return the forms to the advisory instructor. All students who provided a completed consent and assent form were enrolled in the study.

Randomization

All students needing PE credit were cluster-randomized to either PE or yoga immediately prior to the start of the
academic year. Consequently, all students within 9th, 10th, and 11th grades took PE or yoga regardless of their ultimate participation in the study (e.g., providing consent and assent). Randomization was based on English class assignment. There were 10 English classes and all students within a class were assigned to yoga or PE. As a result of scheduling and space requirements, the cluster randomization was performed in a 2:3 ratio with four classes assigned to yoga and six classes assigned to PE.

Instruments

Five instruments measuring mediating variables were used within this study. Two of the subscales of the Response to Stress Questionnaire (RSQ) (Connor-Smith, Compas, Wadsworth, Thomsen, & Saltzman, 2000) were used for analysis whereas composite values were used for the other instruments and therefore the total number of psychosocial variables analyzed was six. The RSQ, Social Stress Version, is a 57-item self-report questionnaire that was used as a measure of emotional regulation processes (Connor-Smith et al., 2000). The RSQ has been shown to have strong internal consistency and test-retest reliability and good criterion validity (Connor-Smith et al., 2000) and has been used in previous work investigating the effects of yoga within a school-based environment (Mendelson et al., 2010). We examined two of the five constructs (24 of the 57 items) within the RSQ because these were most directly theoretically relevant to expected changes because of yoga practice: (1) voluntary engagement, which includes attempts to change the situation or one’s emotions in within the person’s conscious control (nine items: problem solving, emotional control, emotional expression); and (2) involuntary engagement, which involves more unconscious or temperamental reactions (15 items: rumination, intrusive thoughts, physiological arousal, emotional arousal, involuntary action).

The teacher and student versions of the Behavior Rating Inventory of Executive Function (BRIEF) (Psychological Assessment Resources, Lutz, FL, United States) were used to measure executive function. The self-report measure (BRIEF-SR) is composed of 80 items measuring adolescent’s views of their own executive function or self-regulation. Items are rated on a 5-point Likert scale (never, sometimes, often, never true, always true), with higher ratings indicating greater impairment. The BRIEF-SR measures eight nonoverlapping theoretically and empirically derived clinical scales that measure different aspects of executive functioning with the inhibit, shift, emotional control and monitor subscales combining to create a Behavioral Regulation Index (BRI) and the Working Memory, Plan/Organize, Organization of Materials and Task Completion combining to create a Meta Cognition Index. These two indices combine to create a Global Executive Composite (GEC). Only the BRI, MI, and GEC were used within this study for analysis. The BRIEF is psychometrically valid, with adequate internal consistency (α = .80–.98), test-retest reliability (.76–.88), and construct validity established through convergent and discriminant analyses (Gioia et al., 2000; Gioia, Isquith, Retzlaff, & Espy, 2002; LeJeune et al., 2010).

The BRIEF measure completed by the teachers regarding the behavior of the student is composed of 86 items. The eight clinical scales, two indices, and single composite score in the teacher form of the measure are similar to those in the BRIEF-SR described above. As with the self-report form, only the BRI, MI, and GEC were used within this study for analysis. The teacher BRIEF has demonstrated both convergent and divergent validity, as well as high internal consistency (α = 0.8–0.98) and a test-retest reliability of 0.88 (Gioia et al., 2000). For both BRIEF surveys, raw scores were transformed into age- and sex-adjusted T-scores for analysis.

The Child and Adolescent Mindfulness Measure (CAMM) is a 10-item measure assessing mindfulness skills such as present-centered awareness and a nonjudgmental stance toward internal experiences (e.g., “I keep myself busy so I don’t notice my thoughts or feelings”; “It’s hard for me to pay attention to only one thing at a time”). The CAMM is scored on a 5-point Likert scale (never true to always true). Previous results suggest that the CAMM is a developmentally appropriate measure for this population with adequate internal consistency (Greco, Baer, & Smith, 2011).

The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) is a 14-item measure assessing subjective well-being and psychological functioning in which all items address aspects of positive mental health (e.g., “I’ve been feeling optimistic about the future”; “I have been feeling loved”). The WEMWBS is scored on a 5-point Likert scale (none of the time to all of the time). Previous results suggest that WEMWBS is reliable and valid for use in adolescent populations (Clarke et al., 2011; Tennant et al., 2007).

Fidelity of implementation of the intervention was measured via four classroom observations across the academic year using two observers and a standardized form created by the authors based on previous work related to yoga in schools (Gould et al., 2014). The form has 10 items describing general activities considered essential to every class (e.g., arrival routine has been established, the day’s theme is made explicit, and so on), which are marked by the observer as yes/no (occurring, or not occurring). There are an additional seven items, which describe the quality of the teacher–student interactions and one item on the physical environment of the classroom, all of which use positive declarative sentences which are rated by the observer on a 5-point Likert scale from strongly disagree to strongly agree.

GPA was calculated as the numeric average of course scores of all courses taken by the student weighted by credit load of each course using a standard process within New
York City public schools. The GPA from the previous academic year (2013–2014) and the current academic year (2014–2015) were analyzed. Demographic data on race, age, sex, height, weight, and cardiovascular fitness were obtained from the school for analysis. The PE and yoga instructors kept track of class attendance and whether the student actively participated in the class (yes/no decision). Discussions with teachers prior to the study defined failure to actively participate as the student resting on the sidelines and/or not performing class activities. Partial participation, with performance of some activities but not others, was classified as full participation. Therefore, attendance alone without active participation was not counted as participating.

Procedures
The four previously described self-report surveys were collected at three time points (September, February, and May). In addition, the teacher report survey (BRIEF) was collected at mid-October, February, and May. The initial teacher BRIEF survey was delayed 6 weeks from the start of school to allow teachers time to familiarize themselves with the student’s behaviors so they would be able to create a valid BRIEF on each student. At each time point trained research assistants met with students in their respective PE or yoga classes to complete all of the measures in one session lasting approximately 45 min. Blinding to group assignment by testers during self-report survey measurement was not possible within the school environment because for practical purposes measurements needed to occur within each of the yoga and PE program classrooms.

Intervention
Both PE and yoga classes met two times per week for 45 min across the entire academic year (58 scheduled classes). The yoga curriculum (Sonima Foundation, n.d.) used mindfulness and yoga-based exercises with the goal of helping students focus on their work and develop the ability to respond appropriately to various challenging situations. Thematic units were introduced across the entire academic year in the following sequence approximately every 4 weeks: The Power to Connect, The Power of Mindfulness, The Power of the Brain Body Connection, The Power of Integration, The Power to Grow, The Power of Positive Habits. The structure of each class was to begin with a brief seated mental and physical centering exercise (1–3 min) followed by teacher led student discussion focusing on the theme/didactic content of the class for that day time (5–7 min) followed by physical postures and breathing exercises (20–25 min) and a final brief relaxation/meditation (5 min). The yoga curriculum is manualized and in accordance with the New York State and National Physical Education Standards. Descriptive information can be found at Sonima Foundation (www.Sonimafoundation.org).

The two teachers providing the yoga curriculum had a minimum 200-hr general yoga teacher training (Registered Yoga Teacher 200®, Yoga Alliance) and attended 6 days of yoga curriculum training prior to the start of the school year.

Training consisted primarily of instruction based on the contents of the manualized curriculum, practice teaching sessions using role playing, personal stories from existing teachers on common challenges and successful solutions, workshops on social/emotional learning, child/teen development and bullying, and an in-depth study about how to work within school culture. In addition, during the academic year the teachers received bi weekly teacher meetings, professional development workshops (three 2-hr trainings), and a minimum of four classroom observations with feedback. The PE class included weight lifting, stationary biking, fitness exercises such as jumping jacks and push-ups, and common games such as soccer and volleyball. Activities varied arbitrarily based on the teacher’s decisions each week.

Data Analysis
Distributions of scores for all variables were inspected for missing data, normality, and outliers. There were no missing data for the primary outcome of GPA, and there was an average of 9.1% missing data across the four measures (six scales) of psychosocial mediating variables.

The GPA data measured as a percentage were available from four marking periods per year in the 2013–2014 and 2014–2015 school years. Pearson's correlation analysis and intra-class correlation (ICC) analyses were conducted separately for the 2013–2014 and 2014–2015 GPA data to assess the extent to which GPA varied across marking periods within students. There was very little within student variation in grades over the course of the year, the correlations between GPA for any given two marking periods within a year were very high (all above 80%), and the ICC for the four marking periods per student was 87%. Thus, a summary estimate of the average GPA for each student across four marking periods was calculated. Therefore, each student had two GPA values analyzed, one representing the mean GPA from the previous year, and one representing the GPA at the end of the study. The end-of-study measure of GPA is referred to as the mean year GPA and is the primary outcome variable of the study.

Analysis was performed to determine equivalence of groups at baseline of all demographic and dependent variables (Tables 1 and 2). There was a non significant difference in mean year GPA for the previous academic year between assigned groups (2.24 points; \( p = .17 \); Table 1 ). Consequently, in order to control for potential confounding effects of previous year GPA a multivariable linear regression model was fit.
Table 1
Means and Standard Deviation or Percent for Study Variables by Group Condition (PE and Yoga)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Yoga (n = 48)</th>
<th>PE (n = 64)</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Age</td>
<td>15.20</td>
<td>0.94</td>
<td>15.44</td>
</tr>
<tr>
<td>Height</td>
<td>54.94</td>
<td>5.69</td>
<td>57.51</td>
</tr>
<tr>
<td>Weight</td>
<td>129.49</td>
<td>26.59</td>
<td>128.00</td>
</tr>
<tr>
<td>Pacer†</td>
<td>48.00</td>
<td>13.32</td>
<td>52.99</td>
</tr>
<tr>
<td>No. of classes participating</td>
<td>39.32</td>
<td>12.88</td>
<td>52.46</td>
</tr>
<tr>
<td>GPA previous academic year</td>
<td>85.62</td>
<td>7.98</td>
<td>83.38</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Female</td>
<td>29</td>
<td>61.7</td>
<td>25</td>
</tr>
<tr>
<td>Asian</td>
<td>7</td>
<td>14.8</td>
<td>5</td>
</tr>
<tr>
<td>Black</td>
<td>12</td>
<td>25.5</td>
<td>17</td>
</tr>
<tr>
<td>Hispanic</td>
<td>21</td>
<td>44.6</td>
<td>38</td>
</tr>
<tr>
<td>White</td>
<td>7</td>
<td>14.8</td>
<td>5</td>
</tr>
<tr>
<td>Free lunch</td>
<td>31</td>
<td>70.4</td>
<td>40</td>
</tr>
<tr>
<td>Reduced lunch</td>
<td>6</td>
<td>13.6</td>
<td>11</td>
</tr>
<tr>
<td>Full price lunch</td>
<td>7</td>
<td>15.9</td>
<td>8</td>
</tr>
</tbody>
</table>

†FitnessGram measure of cardiovascular health.
*p ≤ .05.

with end-of-study mean year GPA as the dependent variable and variables for yoga versus PE assignment and previous mean year GPA as independent variables (Model 1). Several students (PE N = 10; yoga N = 7) did not attend the school in 2013–2014 and so were dropped from Model 1.

As there was a significant difference in active participation between groups (p = .001; Table 1), a second model (Model 2) was created that added a variable for class participation to Model 1. The measure of participation was for the class the student was assigned to, and therefore if a student switched classes they did not accrue further measured participation. Therefore, in Models 1 and 2, we employed intent to treat approaches to the analysis, and students who dropped out or switched to PE or to yoga were analyzed as if they completed the study as assigned. This is the recommended approach to analyzing randomized trial data. Model 2 includes class participation as a covariate and thus assesses the effect of the assignment to yoga conditional on the extent of participation in the assigned class.

Lastly, it was hypothesized that any benefit of yoga education would only accrue if the student was assigned to yoga classes and participated in the class. Model 3 was fit with an interaction term for class assignment and class participation. The participation variable was centered to the median participation level observed across the pooled group of students assigned to yoga or PE; a participation level of 49 classes. Therefore, Model 3 accounts for crossovers, dropouts and students who attended their assigned class but did not participate. Participation was only counted when students attended and participated in the classes they were assigned to. Thus if a student was assigned to yoga and only participated in 10 classes, his or her participation score was 10. The interaction model thus assesses the association between assignment to yoga and grades at each level of participation in the assigned class. Students who switched classes or went to class but did not engage are thus coded as having low participation in their assigned class. This analysis answers the question: Does the effect of assignment to yoga on grades increase with higher participation in yoga practice?

Scale scores from each questionnaire measuring psychosocial variables and conceptualized as mediating factors were calculated at each of the three time points for each student. To test whether assignment to yoga affected scores on each scale a series of regression models were fit that used the scale score at the end of the year as the dependent variable and variables for class assignment, scale score at the beginning of the year and class participation in yoga or PE as the independent variables.

RESULTS

Recruitment is detailed in Figure 1. There were a total of 283 potentially eligible 9th–11th grade students. Parents of 66 students refused consent, whereas 102 parents failed to respond to the request for consent in any manner despite repeated follow up efforts. Of the 115 students with a parental consent, only three refused assent leaving the total study enrollment of 112 students. There were 64 and 48 participants in the PE and yoga groups, respectively. The difference in group assignment was due to the uneven number of
students within the underlying English classes used for randomization and to the fact that school administrators chose to place two more English classes within the PE program (6 classes) compared to the yoga program (4 classes) because of scheduling and space requirements. The relative percentage of students within those assigned to PE or yoga who enrolled in the study was similar (approximately 38% and 42%, respectively). There was no loss to follow up and all participants originally enrolled were analyzed.

As per Table 1, the groups were not significantly different in age, weight, cardiovascular fitness, race/ethnicity, or socioeconomic status (free lunch status). However, there was a significantly higher percentage of females in yoga compared to PE ($p = .02$), a significantly higher amount of class participation in PE ($p = .001$), and a significantly higher degree of height in participants in PE ($p = .01$). The previous academic year mean GPA for the students assigned to yoga was slightly higher but not significantly different from students assigned to PE. However, because this was the primary outcome variable this variable was controlled for in our analyses as described above. As per Table 2, there were no significant differences in any mediating variables with the single exception of the Involuntary Subscale on the Response to Stress Questionnaire ($p = .39$).

Differences in participation between groups appear to be due to students switching group assignment and students taking double the number of expected classes per week. Ten students in yoga requested and received permission to transfer to the PE program (five in semester one, and five in semester two). Similarly, two students in the PE program transferred to yoga (in semester two). Further, nine students placed in the PE program and one student in the yoga who needed to rapidly acquire PE credits were assigned to four classes per week rather than two classes per week in their respective groups. Consequently, although all students in both groups initially received the allocated intervention, there was a large variation in dosage both within and between groups. Students switching group assignment were analyzed as members of their original group assignment as per intent to treat principles. As described above in the Data Analysis section, Model 3 accounts for crossovers, dropouts, and students who attended their assigned class but did not participate.

Regarding fidelity of the intervention, 98.7% of the 80 measures (10 activities × 4 classes × 2 observers) suggested that the activities considered essential to every class were being implemented. Mean (SD) values of the two observers for the 7 items (5-point Likert scale) which describe the quality of the teacher student interactions was $4.5 \pm .76$ and $4.6 \pm .56$ suggesting that teacher student interactions were being successfully implemented. The kappa for inter-observer agreement for the 7 items was fair ($0.266; p = .070$) (Sim & Wright, 2005).

The end of study 2014–2015 mean (SD) GPA was 85.71 (8.4) and 83.45 (8.79) for the yoga and PE group, respectively. These within-group values and the between-group differences are essentially unchanged from the previous year (cf. baseline values in Table 1). Table 3 presents the results of regression Models 1–3 and shows that previous year GPA significantly predicted mean year GPA in 2014–2015 ($p < .001$), that higher participation was significantly associated with higher mean year GPA ($p = .002$), and that there was a significant interaction between class assignment and participation ($p = .017$). Model 2 estimates that for each class the student participated in, the mean year GPA was 0.08 percentage points higher. Model 3 indicates that mean year GPA

| Table 2 |
|-------------------|-------------------|-------------------|
| **Means and Standard Deviations of Mediating Variables at Three Time Points** |
| **Time 1** | **Time 2** | **Time 3** |
| **n** | **Mean** | **SD** | **n** | **Mean** | **SD** | **n** | **Mean** | **SD** | **p-Value** |
| Self-Report BRIEF GEC | | | | | | | | |
| Yoga | 45 | 53.00 | 14.6 | 38 | 51.66 | 14.8 | 39 | 52.49 | 13.8 |
| PE | 63 | 49.35 | 10.8 | 57 | 49.40 | 12.6 | 55 | 46.84 | 11.9 |
| Teacher Report Brief GEC | | | | | | | | |
| Yoga | 45 | 58.64 | 16.5 | 44 | 61.55 | 18.9 | 45 | 64.51 | 20.4 |
| PE | 64 | 57.48 | 12.0 | 61 | 56.44 | 13.6 | 59 | 59.39 | 16.6 |
| RSQ Voluntary | | | | | | | | |
| Yoga | 46 | 1.35 | 0.6 | 38 | 1.32 | 0.5 | 39 | 1.37 | 0.6 |
| PE | 64 | 1.23 | 0.6 | 60 | 1.35 | 0.6 | 56 | 1.27 | 0.6 |
| RSQ Involuntary | | | | | | | | |
| Yoga | 45 | 0.96 | 0.6 | 38 | 0.90 | 0.7 | 39 | 0.84 | 0.8 |
| PE | 64 | 0.70 | 0.6 | 58 | 0.78 | 0.6 | 55 | 0.67 | 0.6 |
| CAMM | | | | | | | | |
| Yoga | 46 | 22.83 | 8.0 | 37 | 24.76 | 9.4 | 39 | 25.59 | 8.7 |
| PE | 64 | 25.23 | 8.7 | 57 | 24.12 | 9.5 | 53 | 26.38 | 8.5 |
| WEMWS | | | | | | | | |
| Yoga | 46 | 48.26 | 11.0 | 38 | 47.13 | 11.7 | 39 | 48.90 | 13.4 |
| PE | 64 | 49.58 | 13.3 | 60 | 51.72 | 10.0 | 53 | 49.51 | 12.0 |
| **Table 2** Behavioral Regulation Index of Executive Function; GEC = Global Executive Composite; RSQ = Response to Stress Questionnaire; CAMM = Child and Adolescent Mindfulness Measure; WEMWS = Warwick-Edinburgh Mental Well-Being Scale. **p-Value** = group differences yoga versus PE at Time 1; *p < .05.
is significantly higher among students assigned to yoga as compared to PE when participation in class is higher as compared to lower (Table 3; Figure 2). Because the participation variable is centered, the interpretation of the beta coefficient for the class assignment variable is the effect of yoga on mean year GPA for students at the mean level of class participation adjusting for previous year mean year GPA; that is, the difference in grades for students who participated in 49 yoga classes vs. 49 PE classes was significantly higher (2.70 percentage points; \( p = .009 \); Table 3). The effect size at this level of participation was small-to-moderate (Cohen’s \( d = 0.31 \)).

We assessed the potential impact of gender in a series of statistical models. In a univariate model with yoga versus PE assignment predicting mean GPA, assignment to yoga was associated with a 2.78 higher GPA (\( p = .13 \)). When gender was added as covariate to this model, and thus adjusted for, the effect of assignment to yoga diminished to a difference of 2.00 units in mean GPA, and males were found to have a mean GPA that was 3.79 units lower than females (\( p = .04 \)). This suggests that there was some confounding effect of gender on the univariate association between assignment to yoga and mean GPA. However, adding gender to Model 1 (Table 3) did not materially alter the results. That is, once previous year mean GPA is accounted for, gender does not predict mean GPA in the 2014–2015 school year and does not act as a confounder. Similarly adding a variable for gender to Model 2 or Model 3 in Table 3 does not alter the beta coefficients for assignment to yoga or for the interaction term between assignment to yoga and class participation.

Finally, regression analyses were used to test whether yoga versus PE was associated with questionnaire scores at the end of the school year after adjustment for questionnaire scale scores at the beginning of the year and participation in classes. Regression analyses of the six psychosocial
Table 3
Associations Between 2014–2015 Mean Year GPA and Assignment to Yoga Versus PE Classes

<table>
<thead>
<tr>
<th></th>
<th>Model 1&lt;sup&gt;a&lt;/sup&gt;</th>
<th></th>
<th></th>
<th></th>
<th>Model 2&lt;sup&gt;b&lt;/sup&gt;</th>
<th></th>
<th></th>
<th>Model 3&lt;sup&gt;c&lt;/sup&gt;</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta coefficient</td>
<td>(95% CI)</td>
<td>p-value</td>
<td></td>
<td>Beta coefficient</td>
<td>(95% CI)</td>
<td>p-value</td>
<td>Beta coefficient</td>
<td>(95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td>Class assignment: Yoga versus PE</td>
<td>0.58 (−1.25, 2.41)</td>
<td>0.54</td>
<td></td>
<td></td>
<td>1.67 (−0.21, 3.55)</td>
<td>0.08</td>
<td></td>
<td>2.70 (0.69, 4.71)</td>
<td>0.009*</td>
<td></td>
</tr>
<tr>
<td>Previous year GPA</td>
<td>0.98 (0.87, 1.10)</td>
<td>&lt;0.001*</td>
<td></td>
<td></td>
<td>0.98 (0.87, 1.10)</td>
<td>&lt;0.001*</td>
<td></td>
<td>0.95 (0.84, 1.06)</td>
<td>&lt;0.001*</td>
<td></td>
</tr>
<tr>
<td>Participation</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>0.08 (0.03, 0.13)</td>
<td>0.002*</td>
<td></td>
<td>0.05 (−0.01, 0.10)</td>
<td>0.117</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td>0.14 (0.03, 0.25)</td>
<td>0.017*</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Model 1 includes variables for class assignment and mean previous year GPA; presented results are mutually adjusted for the listed variables.

<sup>b</sup>Model 2 includes variables for class assignment, mean previous year GPA, and class participation; presented results are mutually adjusted for the listed variables.

<sup>c</sup>Model 3 includes variables for class assignment, mean previous year GPA, class participation, and an interaction term between class assignment and class participation; presented results are mutually adjusted for the listed variables.

*<i>p < .05.</i>

Fig. 2. Difference in mean GPA 2014–2015 depends on the number of classes participated in.

mediating variable scale scores found no significant associations with group assignment (Figure 3).

**DISCUSSION**

The study’s primary hypothesis that yoga would improve academic performance was not supported by intent to treat analysis. However, it was supported by per protocol analysis that controlled for previous academic year GPA and assessed interactions between assignment and participation. The interaction model estimates that students who participated in 49 yoga classes had significantly improved mean year GPA compared to students who participated in 49 PE classes (mean difference = 2.7 percentage points). Hypotheses regarding the potential mediating effect of emotional regulation and executive function were not supported.

At the median level of participation (49 classes), the mean effect size was 0.312. In a well-respected synthesis of over 800 meta-analyses related to educational achievement, the author (Hattie, 2009) describes 138 potential factors along the unidimensional metric of effect size. Within this synthesis, the value found in this study would be described as small-to-moderate and similar to the level of effects because of homework (0.29), teachers (0.32), and decreasing disruptive behavior (0.34). This value is also comparable to other well-accepted approaches which have been shown to improve academic performance such as social-emotional learning (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011).

The failure of intent to treat analysis to find significant differences in the primary outcome is not surprising given the specifics of this study’s implementation. Intent to treat analysis is a more conservative analytical approach which generally adds confidence regarding effectiveness relative to real-world application. However, the expectation of equivalence at baseline of the primary outcome variables required for a valid interpretation of intent to treat analysis did not occur in this study. In addition, there was large variation between and within groups for dosage of both programs. In effect, some students participated in virtually no classes while others received a large number of classes. Our analysis examining participation as a moderator captures this variability as we would expect dosage to be related to effect. The negative intent to treat findings do reflect the possible real-world effects of trying to deliver a yoga program within a school system. However, the intent to treat approach fails to utilize important information within the data; the student’s actual participation in the classes. Studies in which randomization is more effective in obtaining baseline equivalence of the primary outcome and in which dosage is more consistent across and within groups may produce different results using intent to treat analysis.

Adding confidence to our results is the measurement of fidelity of implementation of the yoga by two observers. Measures indicate that essential activities were being implemented and that the quality of student–teacher interaction was high. The extensive training of teachers as described previously, the manualized curriculum, and the positive measures of fidelity suggest that the curriculum was being delivered as designed.
The findings of this study are in agreement with the two previous studies which found yoga positively influences academic performance (Butzer et al., 2015; Kauts & Sharma, 2009) and in partial agreement with the mixed findings of Smith et al. (2014). These similar findings occurred despite substantial differences in methodology. Kauts and Sharma, (2009) examined seventh and eighth graders and Smith et al. (2014) examined third to fifth graders rather than high school students, and both studies used standardized testing rather than GPA as a primary outcome. This study is most similar to the study of Butzer et al. (2015) in both methodology and findings. Butzer et al. (2015) also examined 9th–11th graders using GPA as the primary outcome variable and a comparison group of PE, but differed in that the intervention was only for 12 weeks as compared to this study’s entire academic year. This study strengthens the existing evidence via its measurement of fidelity and examination of the effects of participation in yoga over an entire academic year. Future study design may benefit from consideration of the results which would have shown no effect had the study been for a single semester.

This study also examined and analyzed potential mediating factors in an attempt to provide preliminary evidence regarding potential mechanisms of the yoga program’s benefits. Changes in self-regulation, executive function, well-being, and mindfulness did not vary by group and consequently did not mediate the effect of yoga on academic performance. These results differ from several exploratory studies within school systems examining high school students which have suggested that yoga can improve anger control and fatigue (Khalsa, Hickey-Schultz, Cohen, Steiner, & Cope, 2012), mood, tension and negative affect (Noggle, Steiner, Minami, & Khalsa, 2012), and emotional regulation (Daly, Haden, Hagins, Papouchis, & Ramirez, 2015). Although the underlying models and psychosocial constructs were generally similar in all of these studies to this study, none used the identical measures of this study, perhaps contributing to the different findings. Also, we note that each of these studies self-identified as preliminary. In one, 37 variables were explored finding two of them significant using a $p$-value threshold of .05 (Khalsa et al., 2012). In the other, 18 variables were explored finding three of them significant using a $p$-value threshold of .025 (Noggle et al., 2012). Only further more rigorous testing will be able to clearly discern the potential effects of yoga on these constructs.

Contrary to our expectations, there was a trend, although nonsignificant, of yoga assignment being associated with poorer scores on the psychosocial scales. These findings are similar to some studies examining elementary school children in which yoga was found to increase negative outcomes such as negative affect (Haden, Daly, & Hagins, 2014), stress perception (White, 2012), or found to be significantly less helpful than PE class in improving self-esteem (Telles, Singh, Bhardwaj, Kumar, & Balkrishna, 2013). Such findings may be related to the nature of self-report measures which only capture self-perception. Increased mindfulness may, in its early stages of practice, increase perceptions of stress (Hayes & Feldman, 2004). The yoga students were specifically and consistently instructed in mindfulness practices as a method to improve self-regulation in the face of environmental stressors. Awareness of stressors, rather than experience of stress, may have affected the self-reports of students.

Although the qualitative results will be reported fully elsewhere, we note briefly here that they are not consistent with the quantitative results. In focus groups at the end of the year, yoga students uniformly expressed an improvement in self-regulation and decreases in mental stress. The difference between qualitative and quantitative results is likely due to either social desirability effects within focus groups, increased awareness of stress as described above (self-report measures), or some combination of these factors. However, these possible explanations cannot account for the teacher
reports of students which also showed no significant difference for students within yoga.

The mechanisms by which yoga may have influenced GPA remains unclear. Certainly the prevailing mechanistic hypotheses within yoga research are that yoga influences self-regulation (Gard et al., 2014). There are several studies in adults (Brown & Gerbarg, 2005; Cappo & Holmes, 1984; Garland, Gaylord, & Fredrickson, 2011; Sauer-Zavalan, Walsh, Eisenlohr-Moul, & Lykins, 2012; Schell, Allolio, & Schonecke, 1994) which support this mechanism. However, most of these studies remain preliminary and lack methodological rigor. In particular, a full understanding of the effects of yoga will require close and prolonged attention to the issue of measurement. If yoga genuinely alters self-regulation, studies would optimally measure not only self-report but also observational behavioral measures and physiological measures (e.g., response to stress, heart rate variability). Such measures were not performed in this study because of the substantial technical, financial, and administrative resources required.

Limitations

There are multiple relevant factors (e.g., student, home, teacher, etc.) contributing to academic outcomes beyond a specific curricular approach (Hattie, 2009). Consequently, the authors suggest that the results of this study examining a new curriculum should be viewed within that wider context of factors. In addition, the interpretation of the positive significant results in this study for yoga should be understood within the study limitations: a relatively small sample size, findings specific to an urban setting, comparison of yoga to a physical activity intervention rather than other proven educational approaches, and grades observed across a single academic year. Larger studies of yoga comparing multiple interventions in diverse settings across multiple years will be needed to provide increased confidence in the outcome.

It is possible that novelty, Pygmalion, or motivational effects may have contributed to the increase in grades of students in the yoga program. As novelty effects are seen before familiarization has occurred, the chance of such effects in this study are minimal given that significant changes in grades occurred months after initiation of the intervention. Pygmalion effects would have required the teachers to become aware of group assignment, to believe that yoga has positive effects beyond that of PE, and to treat yoga program students differently such that it impacted grades. We believe this scenario unlikely, although we acknowledge its possibility. Motivational effects require that students believe their selection to participate conveys special treatment. In this study, all students in Grades 9, 10, and 11 received PE or yoga regardless of study participation. Furthermore, students were aware that assignment of English classes to yoga or PE was performed randomly. Nevertheless, it is possible that students assigned to yoga felt they were receiving special treatment leading to motivational effects which were expressed within their academic classrooms.

This study had several additional limitations including the lack of blinding of tester and an excessive reliance on self-report with its attendant limitations. Analysis is also limited by an inability to test effects due to clustering, which would have been performed by comparing last year’s academic grades of all students in the six English classes which received PE to all the students in the four English classes that received yoga. However, as grades from all students were not available (only grades within those who enrolled in the study were available), this analysis could not be performed.

CONCLUSION

Although intent to treat analyses found no effect of assignment to yoga versus PE on grades, the data support the hypothesis that participation in yoga classes is associated with higher mean GPA when compared to an equivalent amount of participation in PE classes. The mean effect size (0.312) for yoga is small to moderate, but is comparable to other well accepted approaches which have been shown to improve academic performance. Hypotheses regarding the potential mediating effect of emotional regulation and executive function were not supported. Larger studies of yoga comparing multiple interventions in diverse settings across multiple years will be needed to provide increased confidence in the outcome.

Acknowledgments—This work was funded by the Sonima Foundation (www.Sonimafoundation.org).

REFERENCES


Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., & Grimm, K. J. (2009). The contributions of “hot” and “cool” executive function to children’s academic achievement, learning-related


