Effects of a School-Based Garden Program on Academic Performance: A Cluster Randomized Controlled Trial

Jaimie N. Davis, PhD, RD; Katie Nikah, MPH; Matthew J. Landry, PhD, RD; Sarvenaz Vandyousefi, PhD, MS, RD; Reem Ghaddar, MS; Matthew Jeans, MA; Michele Hockett Cooper, MS; Bonnie Martin; Lyndsey Waugh; Shreela V. Sharma, PhD, RD, LD; Alexandra E. van den Berg, PhD, MPH

ABSTRACT

Background School gardening programs have consistently been found to improve dietary behaviors in children. Although several quasi-experimental studies have also reported that school gardens can enhance academic performance, to date, no randomized controlled trial has been conducted to substantiate this.

Objective The objective of the study was to examine the effects of Texas Sprouts (TX Sprouts), a gardening, nutrition, and cooking program vs control on academic performance in primarily low-income, Hispanic children.

Design This is a secondary analysis of the grade-level academic scores from schools that participated in the TX Sprouts program, a school-based cluster randomized controlled trial, consisting of 16 elementary schools that were randomly assigned to either the TX Sprouts intervention (n = 8 schools) or control (delayed intervention; n = 8 schools).

Participants/setting Analysis included 16 schools with students in fourth and fifth grade in Austin, TX from 2016 to 2019 that had a majority Hispanic population and a majority of children participating in the free and reduced lunch program.

Intervention The intervention consisted of 18 one-hour gardening, nutrition, and cooking lessons taught in an outdoor teaching garden by trained educators throughout the academic year.

Main outcome measures Texas Education Agency grade-level data for the State of Texas Assessments of Academic Readiness were obtained via the Texas Education Agency website for the corresponding year of the intervention or control condition.

Statistical analysis performed Repeated measures general linear models with pre- and post-intervention State of Texas Assessments of Academic Readiness scores as the dependent variable were run, adjusting for the percent of free and reduced lunch and school district as covariates.

Results Schools that received the TX Sprouts intervention had a 6.5-percentage-point increase in fourth-grade reading State of Texas Assessments of Academic Readiness scores compared with control schools (P = .047). There were no significant differences in reading scores for fifth grade students or math scores for either fourth- or fifth-grade students between groups.

Conclusions Study findings provide evidence that school gardening programs may have some modest effects on academic achievement.

J Acad Nutr Diet. 2022; - ( - ) - - .

M ANY STATES MANDATE THAT ELEMENTARY schools implement programs to enhance nutrition and child health.1,2 School gardening programs present the ideal setting to address this mandate, as they have consistently been found to increase fruit and vegetable (FV) consumption3-8 and improve dietary-related psychosocial variables, such as preference for FV, willingness to try FV, and self-efficacy to eat FV. In a randomized controlled trial (RCT), an after-school gardening and nutrition program reduced obesity parameters among 319 primarily low-income Hispanic children.2 Texas Sprouts (TX Sprouts) was a cluster RCT to test the effects of an in-school gardening, nutrition, and cooking intervention on dietary intake, obesity, and metabolic outcomes in more than 3,000 third- to fifth-grade students in 16 elementary schools.15
Results showed clinically and statistically significant improvements in vegetable intake, added sugar intake, and glucose control in the TX Sprouts intervention group compared with the control group.\textsuperscript{15,16}

School-based gardening programs can enhance student academic performance in addition to health. Studies have shown that hands-on teaching effectively improves knowledge and enhances learning.\textsuperscript{17,18} More specifically, school gardens have been found to have a positive effect on science, language arts, and math academic achievement scores.\textsuperscript{18} Elementary students learn about the life cycle of plants, soil structure, seed germination, insects and other wildlife, environmental horticulture, and other experiential learning themes. Two quasi-experimental, nonrandomized studies found that a 14-week garden-based intervention resulted in improved science achievement scores compared with traditional classroom teaching in third- to fifth-grade students.\textsuperscript{17,19}

In another nonrandom quasi-experimental study, a multi-level, school-based obesity prevention program that included a school gardening curriculum resulted in improvements in math scores compared with control schools.\textsuperscript{20,21} To date, no RCT has examined the effects of a garden-based intervention on academic performance. Evaluating the effect of school-based garden programs, in addition to the standard educational curricula, on academic performance will help highlight the value of school gardens to education organizations and school administrators.

Therefore, the goal of this ancillary study was to leverage the TX Sprouts cluster RCT to examine the secondary effects of the TX Sprouts intervention vs the control (delayed intervention) on academic performance. It was hypothesized that schools receiving the TX Sprouts intervention would have improved academic scores compared with the control group.

**METHODS**

**Study Design and Participants**

This was a secondary analysis examining the effects of TX Sprouts program on academic outcomes. Although informed consent/assent procedures for participation in primary data collection are described in the TX Sprouts publications,\textsuperscript{14,15} no consent/assent for participation in this secondary analysis of academic performance data was required and a formal Institutional Review Board exemption from The University of Texas was granted. All fourth- and fifth-grade students who attended schools in the TX Sprouts program who completed the State of Texas Assessments of Academic Readiness (STAAR) tests were included in this analysis. The STAAR scores were published on the Texas Education Agency (TEA) website.

The complete design and methodology of the main study has been published previously.\textsuperscript{14,15} The TX Sprouts study was a 1-year school-based cluster RCT with 16 elementary schools from 5 different independent school districts in the Greater Austin area. Of the schools eligible, 20 provided letters of support and the first 16 were randomly assigned to either the intervention (n = 8 schools) or control group (delayed intervention; n = 8 schools). The 4 remaining schools were placed on a contingency list, in case any of the 16 randomly assigned schools dropped out. Of the 16 randomly assigned schools, 2 schools (1 randomized to the intervention and 1 randomized to the control condition) declined to participate due to academic probation because of lower STAAR test scores the year prior. These schools were replaced with the next 2 schools on the contingency list. Inclusion criteria for all schools included high proportion of Hispanic children (>50%); high proportion of children participating in the free and reduced lunch (FRL) program (>50%); location within 60 miles of central Austin; and no existing garden or gardening program. The intervention was implemented over 3 waves, from 2016 to 2019, spanning 1 school year. Randomization was stratified by wave and further details of the randomization process are described elsewhere.\textsuperscript{14}

**Intervention**

With the help of the school stakeholders, families, and volunteers, gardens were built at every intervention school in the spring before the academic year of the implementation of the garden and nutrition lessons. All gardens included 2 raised vegetable beds; 2 in-ground native and herb beds; a large shed for tools and materials; a whiteboard; and seating for approximately 25 students. Schools were provided with the materials and supplies needed for garden upkeep (eg, rakes, hoses, and trowels) and for teaching the lessons (eg, tables, chairs or benches, portable handwashing sink, and pots and pans).

All fourth- and fifth-grade students enrolled in the intervention schools received the TX Sprouts intervention for the full school year. The TX Sprouts curriculum included the following topics: healthy cooking and preparation of FV; making nutritious food choices in different environments; eating locally produced food; low-sugar beverages made with fresh FV; health benefits of FV; how to eat healthfully in food desert neighborhoods; and food equity and community service. Full-time trained nutrition and garden educators taught 18 one-hour TX Sprouts lessons separately in the outdoor garden to each fourth- and fifth-grade class throughout the school year as part of their normal school day. Every lesson included either a garden taste-test or a cooking activity. Every lesson was also mapped on Texas Essential Knowledge and Skills for science, math, language arts, health, and social studies. Dose and reach of the intervention were reported previously.\textsuperscript{15} In brief, <1% of the 18 lessons were modified or shortened across all 8 intervention schools. Although student lessons were designed to be taught outside, approximately
one-third of classes were taught indoors due to inclement weather.

**Control**
The control schools received a delayed intervention (identical intervention as described above) the year after their measurements. Every control school received a garden, identical in size and structure to the intervention schools, and trained educators taught the same in-school garden lessons at the control schools.

**Outcome Measurements**
School size and demographic data were obtained from each school’s website for the academic year that the school was enrolled in the study (intervention or control year, depending on how the school was randomized). Data for STAAR was obtained via the TEA website from the Texas Academic Performance Reports. TEA provides the percent satisfactory scores for grade-level subjects, including reading and math. Because STAAR testing is done at the end of each school year, baseline scores were taken from the previous year for that school, and follow-up scores were used from the year that the school received either the intervention or the control condition. Thus, for fourth-grade students, academic scores from the end of their third-grade year were used for baseline, and scores obtained from the end of their fourth-grade year were used for follow-up, thus capturing fourth-grade students who received either the intervention or control condition their fourth-grade year. Similarly, for fifth-grade students, academic scores from the end of their fourth-grade year were used for baseline and scores from the end of their fifth-grade year were used for follow-up, thus capturing fifth-grade students who received the intervention or control condition during their fifth-grade year. Of note, baseline STAAR data are not available for writing and science scores, therefore, those scores are not included in this analyses. Special education children and/or children with learning disabilities are not required to take STAAR testing, therefore, their data were not included in this analysis.

**Statistical Analyses**
Statistical analyses were completed using SPSS software, version 26. Categorical data were summarized as counts and percentages. Baseline characteristics including sex, race and ethnicity, and FRL eligibility status were compared between intervention and control groups using independent t tests. Univariate general linear models were run to assess differences in mean STAAR scores at baseline between intervention and control schools. General linear models repeated measures were used to test differences between intervention and control STAAR scores from pre to post intervention, with the mean and SE and schools treated as random clusters. Of note, STAAR data were only available at the grade level per school, as a result within-student changes across the school year could not be examined. Hence, the data analyses consisted of grade-level summary statistics for baseline and 1-year calculated from grade-level data. These

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Combined (16 schools)</th>
<th>Intervention (8 schools)</th>
<th>Control (8 schools)</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total fourth- and fifth-grade students enrolled</td>
<td>3,114</td>
<td>1,400</td>
<td>1,714</td>
<td>—</td>
</tr>
<tr>
<td>Fourth grade</td>
<td>1,554</td>
<td>699</td>
<td>855</td>
<td>—</td>
</tr>
<tr>
<td>Fifth grade</td>
<td>1,560</td>
<td>701</td>
<td>859</td>
<td>—</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td>.12</td>
</tr>
<tr>
<td>Male</td>
<td>1,601 (51.4)</td>
<td>731 (52.2)</td>
<td>868 (50.6)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1,513 (48.6)</td>
<td>669 (47.8)</td>
<td>847 (49.4)</td>
<td></td>
</tr>
<tr>
<td>Race and ethnicity</td>
<td></td>
<td></td>
<td></td>
<td>.69</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>137 (4.4)</td>
<td>63 (4.4)</td>
<td>74 (4.3)</td>
<td>.96</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>2,177 (69.9)</td>
<td>972 (69.4)</td>
<td>1,207 (70.4)</td>
<td>.90</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>498 (16)</td>
<td>234 (16.7)</td>
<td>264 (15.4)</td>
<td>.86</td>
</tr>
<tr>
<td>Non-Hispanic Black</td>
<td>302 (9.7)</td>
<td>133 (9.5)</td>
<td>170 (9.9)</td>
<td>.52</td>
</tr>
<tr>
<td>Free and reduced lunch status</td>
<td></td>
<td></td>
<td></td>
<td>.69</td>
</tr>
<tr>
<td>Eligible</td>
<td>2,398 (77.0)</td>
<td>1,057 (75.5)</td>
<td>1,337 (78.0)</td>
<td></td>
</tr>
<tr>
<td>Ineligible</td>
<td>716 (23.0)</td>
<td>343 (24.5)</td>
<td>377 (22.0)</td>
<td></td>
</tr>
</tbody>
</table>

*P value is from independent t tests between intervention and control schools.
summary statistics were treated as repeated measures with schools as the unit of analysis. A priori covariates for all models included school district and %FRL. Statistical significance was set at $P < .05$.

RESULTS

The STAAR test scores from 3,114 eligible students from the 16 schools were included in this analysis. Table 1 displays the demographic characteristics for students in the intervention vs control schools. The schools served primarily Hispanic (73%) and FRL-eligible (77%) populations. There were no differences in sex, race and ethnicity, or FRL eligibility between control and intervention schools.

Table 2 displays the baseline and change in STAAR scores between intervention and control schools. Univariate GLM showed no differences in any of the baseline STAAR scores between intervention and control schools. Using GLM repeated measures, schools that received the TX Sprouts intervention had a 6.5-percentage-point increase in fourth-grade reading STAAR scores compared with control schools ($P = .047$). There were no statistically significant differences in changes in fourth-grade math scores or fifth-grade reading or math scores between intervention and control groups.

DISCUSSION

The current study is the first RCT to examine how a gardening, nutrition, and cooking intervention impacts academic achievement scores, in elementary schools that serve primarily Hispanic and low-income populations. These findings support that school-based gardening, nutrition and cooking programs improve reading scores in fourth-grade students. These findings will potentially help schools and school districts make the case for additional funding, support, and training for school garden program implementation.

There are many possible mechanisms to consider when trying to understand how a garden-based intervention would enhance academic performance. One possible explanation is that TX Sprouts intervention resulted in numerous health benefits, such as improved vegetable intake and reduced sugar intake, all of which could aid in helping the child perform better academically. Studies have shown that eating more fruits and vegetables is associated with increased academic performance in children. Other studies have shown that lower added sugar intake is linked to improved learning and cognition. These studies support the notion that healthier children learn better.

Each garden lesson for the TX Sprouts program was mapped on the Texas Essential Knowledge and Skills for science, math, language arts, health, and social studies. Given that most states mandate that schools provide nutrition education in schools, using a garden and nutrition–based curricula helps schools meet those nutrition mandates, while also helping improve other subjects. To date, most garden-based interventions showing an improvement in academic performance used horticultural-based curricula, like Junior Master Gardener, in which the main focus is not on nutrition. Many studies that found positive associations between garden-based learning and math and science scores varied widely in the curricula that were connected to the gardens. For example, plant taxonomy, plant parts, flower dissection, soil chemistry, insects and other wildlife, and ecology and

<table>
<thead>
<tr>
<th>Variable</th>
<th>STAAR Scores (%)</th>
<th>Differences in mean percentage</th>
<th>Intervention effect</th>
<th>95% CI for difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intervention</td>
<td>Control</td>
<td>$P$ value</td>
<td>$P$ value</td>
</tr>
<tr>
<td></td>
<td>Baseline (SE)</td>
<td>Baseline (SE)</td>
<td>$P$ value for baseline</td>
<td>$P$ value for baseline</td>
</tr>
<tr>
<td></td>
<td>Baseline (SE)</td>
<td>Baseline (SE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fourth-grade reading</td>
<td>63.13 (4.62)</td>
<td>68.38 (3.84)</td>
<td>.355</td>
<td>6.5 (4.14)</td>
</tr>
<tr>
<td>Fourth-grade math</td>
<td>65.38 (3.56)</td>
<td>67.13 (5.13)</td>
<td>.358</td>
<td>1.5 (4.91)</td>
</tr>
<tr>
<td>Fifth-grade reading</td>
<td>75.38 (4.21)</td>
<td>77.13 (4.06)</td>
<td>.288</td>
<td>0.88 (2.89)</td>
</tr>
<tr>
<td>Fifth-grade math</td>
<td>83.00 (2.90)</td>
<td>82.88 (3.75)</td>
<td>.306</td>
<td>1.13 (2.85)</td>
</tr>
</tbody>
</table>

$^{a}$STAAR = State of Texas Assessments of Academic Readiness.
$^{b}$STAAR scores reflect percent of students that met satisfactory status.
$^{c}$Analyses of variance was run to determine differences in baseline STAAR scores between intervention and control groups.
$^{d}$Differences in the mean percentage between intervention and control and SE of the differences in the mean percentage of each change.
$^{e}$Repeated measures general linear modeling with the pre- and post-intervention STAAR scores as the dependent variables were run, adjusting for percent free and reduced lunch and school district.
environmental horticulture, to name a few, were central science themes presented in studies that found a link between garden-based learning and improvements in math and science scores.18 The TX Sprouts curriculum improves both health outcomes15 and reading scores in fourth-grade students. This is likely because every TX Sprouts lesson had a reading and writing component and all children were given garden journals, which they used for journaling during and outside of the scheduled TX Sprouts classes.

Schools and school districts are often reluctant to add additional programming that does not directly improve standardized testing. In the current study, 2 of the elementary schools that agreed to participate in the program and were randomized decided to drop out before any recruitment could occur due to being placed on academic probation because of low state test scores the year before. The principals informed the study team that any program that could potentially take away from preparation for standardized tests would not be allowed. In a cross-sectional study of 109 school gardens in the Greater Austin area, not having adequate administrative and/or district support were the top barriers to a thriving school garden program.32 The current findings show that a garden-based program added to existing school curricula can enhance standardized test scores, which may help to encourage other administrators and school district personnel to adopt and/or support garden-based learning.

Limitations
One limitation is the small sample size when the analyses is conducted at the school level—8 intervention schools vs 8 control schools. This may increase the likelihood of type 2 error. Another limitation is that this secondary analysis was limited to fourth- and fifth-grade students, and the TX Sprouts intervention was delivered to all third- through fifth-grade students in participating schools. However, given that the baseline STAAR scores were obtained for the prior year and compared with the follow-up year after exposure to either the TX Sprouts intervention or the control condition, the third-grade students would not have had baseline data and were therefore excluded from the grade-level analyses. Similarly, writing and science STAAR scores were not analyzed, as writing is only tested in fourth-grade and science is only tested in fifth-grade. Another limitation is that the TEA website does not provide grade-level overall scores, just grade-level subject scores. Nonetheless, fourth and fifth grades are key years for academic tests. Finally, another limitation is that no corrections were made for multiple comparisons.

CONCLUSIONS
This study provides evidence that school gardens may result in modest improvements in reading levels in schools, in addition to other analyses from the TX Sprouts program that showed enhanced health of children. Although the current evidence can be used to encourage school stakeholders, such as teachers, administrators, and school district personnel, to support garden-based learning in classrooms and schools, more research is warranted to assess the effects of school gardening programs on aggregate school test scores and/or overall grades, as well as other subjects, such as writing and science.

References


