STEMulating Interest: A Meta-Analysis of the Effects of Out-of-School Time on Student STEM Interest

Jamaal Young¹, Nickolaus Ortiz², Jemimah Young³
¹University of North Texas
²Texas A&M University
³University of North Texas

To cite this article:

This article may be used for research, teaching, and private study purposes.

Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

Authors alone are responsible for the contents of their articles. The journal owns the copyright of the articles.

The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material.
STEMulating Interest: A Meta-Analysis of the Effects of Out-of-School Time on Student STEM Interest

Jamaal Young, Nickolaus Ortiz, Jemimah Young

Abstract

This study is a meta-analysis of the effects of out-of-school time (after school, summer camps, enrichment programs, etc.) on the student interest in STEM. This study was guided by the following research questions: (1) How effective is OST as a means to foster student interest in STEM? (2) How does the effectiveness of OST differ by program and study characteristics? A total of 19 independents effect sizes were extracted from 15 studies investigating the effect of out-of-school time (OST) on STEM interest. Included studies were representative of K-12 settings in the United States from 2009-2015. Specifically studies were included if they directly assessed the effects of OST on STEM interest, and provided sufficient data to calculate an effect size. The status of publication was not a constraint on this investigation, thus grey literature was included along with journal articles to provide a more representative sample of studies. The results suggest that out-of-school time has a positive effect on student interest in STEM. Furthermore, the variation in these effects is moderated by program focus, grade level, and the quality of the research design. The effects of out-of-school time on STEM interest are synthesized, and implications for teaching and practice are provided.

Introduction

Promoting and sustaining interest in Science Technology Engineering and Mathematics (STEM) is a major concern for the future competitiveness of the United States in a global economy. Despite recent increases in the number of college-educated U.S. citizens, many jobs in STEM fields continue to go unfilled due to a lack of quality applicants (Atkinson, 2013). Given these concerns, the United States considers the enrollment and retention of students in STEM-related majors a high priority (Committee on Science, Engineering, and Public Policy, 2007). Many factors contribute to a student’s decision to declare a STEM major and subsequently enter a STEM profession. Proficiency and interest are recognized amongst these factors as the two most influential in student progression through the STEM pipeline (Beir & Rittmayer, 2008; Business-Higher Education Forum, 2011). Historically, policies have focused on the development of knowledge and skills in the area of STEM. However, interest in STEM is declining in high school students for most student populations, with the exception of Asian students (Le, & Gardner, 2010). Appropriately, programs have shifted from a singular focus on STEM achievement to a more holistic approach to STEM talent development. For example, the Educate to Innovate program launched by the Executive Office of the President (2009) was developed to address the persistence of the achievement gap, with a secondary aim to increase STEM interest in traditionally underrepresented populations, such as women and students of color. In recognition of the need to spark STEM interest among K-12 youth through the development of quality engineering experiences, there is widespread support for STEM-focused outreach programs and informal learning opportunities (Decoito, 2014; Frantz et al., 2011; Jordan & Young, 2010). These outreach programs and informal learning opportunities are typically situated in out-of-school time (OST) activities.

OST activities provide unique opportunities to develop future STEM talent by means beyond the capacity of the allotted school time. OST programs provide expanded learning, which includes a plethora of content-rich opportunities outside of school time (Bevan & Michalchik, 2013). In addition, OST STEM activities allow students to meet STEM professionals and learn about STEM careers (Fadigan & Hammrich, 2004). These programs help learners to expand their identities as achievers in the context of STEM (Barton & Tan, 2010). Thus, as learners become actively involved in producing scientific culture, they come to recognize the utility of STEM as it relates to their personal lives (Bell et al., 2012). As described above, when done well STEM OST
programs engage students in rigorous, high-quality, and purposeful activities (Gupta et al., 2011). Yet, although STEM OST programs have increased substantially (Bell et al., 2009) summaries of their effectiveness as means to foster STEM interest remain elusive (Adams et al., 2008; Guiterrez, 2012). Although research of the factors affecting STEM interest are still emerging (Banning & Folkestead, 2012), an examination of the effects of OST on STEM interest is necessary given the proliferation of STEM OST programs. Thus, we conducted this synthesis to address the following research questions:

1. How effective is OST as a means to foster student interest in STEM?
2. How does the effectiveness of OST differ by program and study characteristics?

Out-of-School Time and Student Interest in STEM

Students of all ages are required to attend some form of schooling, and this ubiquitous requirement consequently allows school to account for noteworthy proportions of a student’s conscious hours. Researchers have distinguished time spent at school in what can be thought of as the instructional period and the remainder of the time OST. Laurer et al. (2006) defined OST as the allotted time where children “are doing something other than activities mandated by school attendance” (p. 276), and stated that this occurs most often in after-school programs and during the summers. Preferential options that afford students an opportunity to pursue their personal interest offer a greater impact and will compose most of their OST; these options often play a role in the success of students during the instructional period. This preferential aspect of OST programs is the impetus to the recent proliferation of STEM related activities to foster interest and achievement in STEM subject matter.

STEM-related OST activities are designed to either increase achievement in STEM content, foster interest in STEM, or a combination of these outcomes. A report by the Afterschool Alliance (2011) asserted that students who attend STEM programs in their OST are better prepared in areas such as problem solving, critical thinking, and collaborating with others. These skills are necessary as well as beneficial because they translate to environments that are not exclusively mathematics-based. While, according to Cooper & Heaverlo (2013) exposure to STEM OST programs has the capacity to generate positive attitudes towards STEM content. STEM OST programs have three consistent outcomes across administrations and designs: 1) increased interest in STEM and STEM learning activities 2) students develop a capacity to productively engage in those STEM learning activities, and 3) students come to appreciate the goals of STEM and STEM learning activities (Krishnamurthi et al., 2011). Despite consistent evidence of the benefits of STEM OST programs, more research is needed to make generalizable decisions concerning the factors that differentiate the success of these programs. The two most common formats for STEM OST programs are summer enrichment programs or after school programs.

Summer Enrichment Programs

Summer schools were created as a way to help students who performed below level to receive supplemental instruction that did not detract from the routine learning schedule during the school year (Smink, 2012). Because summer school has the potential to help students continue to build academic skills when most students are not enrolled in academic programs, many summer schools evolved into summer enrichment programs to support accelerated learning and early graduation (Berliner, 2009). These accelerated learning opportunities are extremely beneficial for girls, students of color, and English language learners. Thus, a growing number of summer enrichment programs are designed to support the growing needs of culturally and linguistically diverse students and students living in poverty (Keiler, 2011; Matthews & Mellom, 2012). Many colleges and universities host STEM summer camps each summer aimed at developing interest and achievement in K-12 learners.

The Aggie STEM center host enrichment camps that last for 13 days aimed at increases STEM interest and achievement. The results of a follow up study suggest that participation in the STEM summer enrichment camp was related to SAT performance and college matriculation (Boedeker, Bicer, Capraro, Capraro, Morgan, & Barroso, 2015). Evidence that summer enrichment programs affect the academic success of diverse students supports the need to better ascertain their influence on attitudes towards STEM. Aside from summer enrichment programs, after school programs are another popular OST learning opportunity. Summer schools have a focus that is more oriented to the need for student improvement in academics, while after-school programs may have a variety of expected goals and outcomes. Given these nuances a better understanding of the comparative effectiveness of these programs necessary.
After-School Programs

Students who participate in activities after school will have less time to become immersed in activities that lead to maladaptive behaviors, such as crime, drug abuse, violence, or sex (Hirsch, Mekinda, & Stawicki, 2010). The range of after-school programs is wide and thus includes a variety of goals, missions, and focuses. Dryfoos (1999) described some categories within after-school programs such as the school-administered and the community-based organization-administered. The school-administered programs include extracurricular activities and extended day programs that focus on academic achievement, providing safe havens for students, and recreation. Community-based organization-administered programs include non-profit agencies that help children to develop academic skills and prevent high-risk behaviors. Because older students do not wish to feel like they are being ‘babysat’ in most cases, and require more opportunities to build healthy attitudes and self-sufficiency, the developmental appropriateness of after school programs is consistently cited as a mediating factor related to program success (Harms, 2013; Webster, Monsma, & Erwin, 2010). Thus, after school programs must be appealing and appropriate for age groups, while providing creative programming and enrichment activities to sustain participation (Grossman, J. B., Walker, K., & Raley, R. (2011). For example, Harmony Public Schools utilize after school programs to further their mission to support STEM learning. Specially, Harmony Public Schools host several STEM related after school programs: (a) robotics, MATHCOUNTS, American Mathematics Contest (AMC), Science Olympiad, and University Interscholastic League (UIL) to name a few ( Sahin, Ayar, & Adiguzel, 2014). Participating in afterschool programs, increases notions of self-perception, reduces maladaptive behaviors, and promotes growth in assessment measures, yet summaries of the effects of after school program effects on STEM attitudes remain elusive. STEM after school and summer enrichment programs are relatively new compared to traditional after school programs and summer enrichment activities; thus it is imperative that the selection of appropriate moderators is guided by sound theoretical and practical knowledge.

Research on Moderators of STEM Interest

Based on prior research related to OST and STEM interest, we identified the following program characteristics as possible moderators of the effects of OST on STEM interest: (1) time frame, (2) program focus, and (3) grade level. Time frame refers to whether the OST program was delivered to students after school, during the summer, or in another time-related format. This format was chosen because it aligns with the designs of previous research on the effects of OST (Cooper et al., 2000; Lauer et al., 2006). Given that the duration of each program is highly contingent upon the time frame, duration was not investigated as a moderator in this investigation. For example, summer programs are constrained by the summer months, while after school programs are typically aligned to the academic year. However, OST program content vary in focus programs, hence the focus of the OST program was identified as a second moderator of STEM interest.

STEM OST programs typically have two foci: academic or social. In the Lauer et al. (2006) study, OST program focus was a significant moderator of mathematics achievement effects in at-risk youth. Given that OST program focus was identified as a moderator of mathematics achievement and the relationship between interest and achievement, it is important to determine how the program focus affects student interest in STEM. Finally, prior research suggests that OST programs effects are differentiated across elementary, middle, and secondary grade levels (Cooper et al., 2000; Grossman, Walker & Raley, 2011). These prior studies examined the effects of OST on student achievement; nonetheless their results remain relevant to the current investigation given the significant relationship between interest and attitudes towards STEM and student performance (Choi & Chang, 2009). Demographics represent another characteristic of the studies pertinent to the measurement of the effect sizes.

Gender, ethnicity, and socioeconomic status (SES) are widely recognized indicators of STEM interest, when measured by one’s choice of college major (Porter & Umbach, 2006). Thus, these variables are considered appropriate moderators of the effects of OST on STEM interest. According to McGrayne (2005), students often see STEM as predominately white, male, and middle class. National statistics mirror these perceptions, with African Americans, Native Americans, and Latinos between the ages of 18–24 accounting for 34% of the population in this age category, yet earning only 12% of the undergraduate engineering degrees (Bean et al., 2014). Additionally, the fact that women remain underrepresented in STEM fields is hard to dispute (Blickenstaff, 2005). Retention data for most STEM fields show a pattern of decreasing representation of women as the level of education increases (Goulden et al., 2011). Despite these trends in enrollment and completion of STEM degrees, trends in STEM interest may tell a different story. While white males are
traditionally well-represented in STEM fields, they historically have the lowest level of interest in scientific fields (Elliot et al., 1996). Contrarily, according to the Business-Higher Education Forum (2011), students of color are more likely to be interested in STEM, but tend to lack the mathematics proficiency. Thus, patterns in student interest in STEM do not mirror patterns in student achievement between racial subgroups and warrant investigation as moderators of STEM interest.

Method

We conducted separate searches using the keywords “Summer Camps” (17 citations), “After school programs” (24 citations), “STEM interest” (33 citations), and “STEM career interest” (10 Citations). Each search was conducted in the following databases: (a) Academic Search Complete, (b) PsycINFO, (c) ERIC, and (d) Dissertation Abstracts. Our search was exhaustive, thus publication date restrictions were not employed. The three searches resulted in 84 citations, which were entered into a master library using Zotero online software. We used Boolean operators to identify studies that incorporated a combination of pertinent search terms. For example, studies that investigated “summer camps” and “STEM interest” were located from within the master list. As a result, the total number of articles that we organized and read was 84. After applying the inclusion criteria a final pool of 15 studies representing 20 independent effect sizes.

Inclusion Criteria

We used the following criteria for including studies:

1. Studies had to concern an OST program for K-12 students. We defined OST programs as an education intervention delivered outside of the regular school day that included but was not limited to summer camps, after school programs, or academic fairs.
2. Studies had to directly assess students’ STEM interest. Examples include survey results, course enrollments, or observational methods.
3. Studies had to disaggregate student results for specific OST programs. For instance, one study included science fairs and science clubs, which represent two separate interventions.
4. Studies could be published or unpublished, including reports, conference proceedings, and dissertations.
5. Studies had to include sufficient quantitative information to calculate effect sizes.

Coding Studies

Each study was coded for information about the OST program characteristics, student sample, and research quality. Program characteristics included time frame (summer camp, after school, special event/experience). Although duration is a reasonable study characteristic it was not included because the time frame correlates strongly with the duration, given after school programs are during the school year and summer camps are shorter summer durations. The primary focus of the OST program is another variable that could influence student interest in STEM. According to Heilbronner (2011) students need to have positive attitudes towards STEM in order to develop interest in STEM careers. Furthermore, developing strong student interest in STEM requires educators to look beyond achievement and tap into student curiosity through socially constructed learning opportunities that are enjoyable (Kapp et al., 1992). Hence, the focus of the OST STEM activity can potentially alter student interest in STEM. Student information included primarily study demographics such as gender representation (male, female, & mix), racial composition based on representation (representative or unrepresentative), and grade level (K–5, 6–8, and 9–12). When grade levels overlapped categories, we chose to categorize studies based on the grade level that was most represented in the sample. Each author met to develop the coding protocol, developed the coding form, and came to a consensus on the overall coding procedure. Following the initial meeting, each author separately coded a random sample of four studies using the coding form. Given their backgrounds and expertise, coding forms from authors 2 and 3 were used to assess inter-rater reliability. The resulting inter-rater agreement was 89.6% (Cohen’s k = .792, p < .001). We compared completed forms, identified and resolved discrepancies, and made appropriate revision to improve performance. The first author reviewed the studies independently of the author pairs and verified the accuracy of the study codes entered in the meta-analysis database.

Study quality was coded as either low, medium, or high based on the studies’ ability to address construct, internal, external, and statistical validity (Valentine & Cooper, 2003). To standardize this process studies were
assessed on each of the aforementioned types of validity (1 = unclear, 2 = somewhat clear, 3 = clear). The total score out of 12 indicated the methodological quality; the higher the score, the better the methodological quality. Given their methodological backgrounds and expertise the first and third authors coded a random sample of five studies to assess inter-rater reliability of the methodological quality. The resulting inter-rater agreement was 87.9% (Cohen’s κ = .776, p < .001). The second author was then consulted to help reconcile conflicts.

Analysis

We conducted the meta-analysis in four steps. First we computed an effect size for each study. Second we computed an overall effect size across the research studies. Then we performed the homogeneity analysis, followed by the final moderator analysis. We utilized Comprehensive Meta-analysis (CMA) version 2.0 for the data analysis and presentation of the results. For the purpose of this analysis, we report Hedges g as the measure of effect sizes, which was calculated and adjusted for small sample sizes within CMA 2.0 (Rosenthal, 1991). There was variation in the design and presentation of study results. For example, some studies presented pretest and posttest scores, while others only presented gain scores. Accordingly, for all studies we used the pooled standard deviation to account for the different standard errors and sample sizes (Hedges & Olkin, 2014). Some studies report outcomes for independent samples on separate intervention, which were analyzed as independent samples.

Data from independent samples were used to compute overall effect sizes for the effects of OST on STEM interest. Based on the assumption that larger sample sizes produce more reliable estimates of effects, studies were weighted according to sample size. We conducted a homogeneity analysis to determine whether the effect sizes varied more than what are expected from sampling error. The value of the Q statistic was statistically significant; thus we concluded that the effect sizes were not homogeneous. This result is consistent with prior research that suggests that STEM interest is differentiated by student and OST program characteristics. Thus, the random effects model was employed and the final moderator analysis was conducted to identify factors that might account for variation in effect sizes across studies. According to Pigott (2012), a random effects moderator analysis is best suited for investigations of multiple sources of variation amongst studies that can be accounted for by study characteristics. Therefore, given the limited set of categorical moderator variable identified in this study and our focus on the study characteristics, the random effects model was used to calculate a Q statistic for each moderator.

Results and Discussion

Table 1 presents the characteristics of the 15 studies included in the synthesis. The publication years for the studies ranged from 2009–2015, and the median year of publication was 2013.

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Student grade level</th>
<th>Representation</th>
<th>Gender</th>
<th>Program Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dabney et al. (2012)</td>
<td>6,882</td>
<td>High School</td>
<td>Unrepresentative</td>
<td>Mixed</td>
<td>Science Club</td>
</tr>
<tr>
<td>Demetry et al. (2009)</td>
<td>176</td>
<td>High School</td>
<td>Unrepresentative</td>
<td>Female</td>
<td>Summer Camp</td>
</tr>
<tr>
<td>Heaverlo (2011)</td>
<td>871</td>
<td>Middle School</td>
<td>Unrepresentative</td>
<td>Female</td>
<td>After School</td>
</tr>
<tr>
<td>Hughes et al. (2013)</td>
<td>27</td>
<td>Elementary</td>
<td>Representative</td>
<td>Mixed</td>
<td>Summer Camp</td>
</tr>
<tr>
<td>Innes et al. (2012)</td>
<td>227</td>
<td>Elementary</td>
<td>Unrepresentative</td>
<td>Mixed</td>
<td>Field Trip</td>
</tr>
<tr>
<td>Knezek et al. (2013)</td>
<td>246</td>
<td>Middle School</td>
<td>Unrepresentative</td>
<td>Mixed</td>
<td>After School</td>
</tr>
<tr>
<td>Kong et al. (2014)</td>
<td>1580</td>
<td>Middle School</td>
<td>Representative</td>
<td>Mixed</td>
<td>Summer Camp</td>
</tr>
<tr>
<td>Mayberry (2015)</td>
<td>31</td>
<td>Middle School</td>
<td>Representative</td>
<td>Female</td>
<td>Girl Scout STEM</td>
</tr>
<tr>
<td>Mohr-Schroeder et al. (2014)</td>
<td>144</td>
<td>Middle School</td>
<td>Unrepresentative</td>
<td>Mixed</td>
<td>Summer Camp</td>
</tr>
<tr>
<td>Nugent et al. (2012)</td>
<td>379</td>
<td>Middle School</td>
<td>Representative</td>
<td>Mixed</td>
<td>Camp and Competition</td>
</tr>
<tr>
<td>Sahin (2013)</td>
<td>379</td>
<td>High School</td>
<td>Representative</td>
<td>Mixed</td>
<td>Science Club</td>
</tr>
<tr>
<td>Schmidt (2014)</td>
<td>49</td>
<td>Middle School</td>
<td>Representative</td>
<td>Mixed</td>
<td>Science Olympiad</td>
</tr>
<tr>
<td>Stoeger et al. (2013)</td>
<td>208</td>
<td>Elementary</td>
<td>Unrepresentative</td>
<td>Female</td>
<td>Mentoring Program</td>
</tr>
<tr>
<td>Tyler-Wood et al. (2012)</td>
<td>32</td>
<td>Elementary</td>
<td>Unrepresentative</td>
<td>Female</td>
<td>Hands of Problem Solving</td>
</tr>
<tr>
<td>Wyss et al. (2012)</td>
<td>72</td>
<td>Middle School</td>
<td>Unrepresentative</td>
<td>Mixed</td>
<td>Video Interview</td>
</tr>
</tbody>
</table>
The majority of the studies were conducted with students in grades 6th–8th or middle schools. The majority of the studies included non-representative samples of students of color. Furthermore, the studies in this sample included mixed gender groups or exclusively female participants. The sample of studies was comprised of studies conducted in the United States, however this was not an inclusion criteria. Finally, the OST activities varied from summer camps to mentoring programs.

We calculated effect sizes for each of 20 independent samples yielded from 15 studies. Table 2 presents information on each independent sample, the treatment students, OST activity, effect size, and lower and upper limits of the 95% confidence interval. Based on the test for homogeneity we concluded that their was significant heterogeneity, $Q(18) = 203.4, p < 0.0001$. The “one study removed” procedure was utilized to identify possible outliers (Borenstein, Hedges, Higgins, & Rothstein, 2009). This approach revealed that one study was located outside of the 95% confidence interval for the average effect size, and thus this study was removed.

After removing the Kong et al. (2014), the overall effect size was determined to be $0.37, p<0.0001$ based on the random effects model. To assess the stability of the summary effect size we calculated the classic fail-safe $N$. According to Rosenthal (1979) the Fail Safe $N$, estimates the number of studies required to yield a non-statistically significant mean effect size at the $p<0.05$ level. Hence, this statistic “indicates the stability of meta-analytic results when additional findings are included, no matter the source” (Persuad, p. 125, 1996). For the present study the value of the Fail Safe $N$ was 1,033, which suggest that we would need to retrieve an additional 1,033 studies to observe a statistically non-significant mean effect size at the $p<0.05$ level.

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample size</th>
<th>Grade level</th>
<th>Effect size</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dabney et al. (2012)$a$</td>
<td>3,602</td>
<td>$12^{th}$</td>
<td>.83</td>
<td>.74 .92</td>
</tr>
<tr>
<td>Dabney et al. (2012)$b$</td>
<td>3,280</td>
<td>$12^{th}$</td>
<td>.73</td>
<td>.65 .82</td>
</tr>
<tr>
<td>Demetry et al. (2009)</td>
<td>176</td>
<td>$7^{th}$</td>
<td>.35</td>
<td>-.02 .72</td>
</tr>
<tr>
<td>Heaverlo (2011)</td>
<td>871</td>
<td>$6^{th}$–$12^{th}$</td>
<td>.26</td>
<td>.13 .40</td>
</tr>
<tr>
<td>Hughes et al. (2013)$a$</td>
<td>13</td>
<td>$5^{th}$–$9^{th}$</td>
<td>.43</td>
<td>-.06 .93</td>
</tr>
<tr>
<td>Hughes et al. (2013)$b$</td>
<td>14</td>
<td>$5^{th}$–$9^{th}$</td>
<td>-.01</td>
<td>-.54 .53</td>
</tr>
<tr>
<td>Innes et al. (2012)</td>
<td>227</td>
<td>$4^{th}$–$9^{th}$</td>
<td>.17</td>
<td>.01 .33</td>
</tr>
<tr>
<td>Knezek et al. (2013)</td>
<td>246</td>
<td>$7^{th}$</td>
<td>.12</td>
<td>-.10 .35</td>
</tr>
<tr>
<td>Kong et al. (2014)</td>
<td>1580</td>
<td>$7^{th}$–$8^{th}$</td>
<td>1.96</td>
<td>1.88 2.05</td>
</tr>
<tr>
<td>Mayberry (2015)</td>
<td>31</td>
<td>$7^{th}$–$8^{th}$</td>
<td>.67</td>
<td>.20 1.54</td>
</tr>
<tr>
<td>Mohr-Schroeder et al. (2014)</td>
<td>144</td>
<td>$7^{th}$–$8^{th}$</td>
<td>.59</td>
<td>.30 .88</td>
</tr>
<tr>
<td>Nugent et al. (2012)$a$</td>
<td>307</td>
<td>$7^{th}$–$8^{th}$</td>
<td>0.1</td>
<td>-.06 0.26</td>
</tr>
<tr>
<td>Nugent et al. (2012)$b$</td>
<td>72</td>
<td>$7^{th}$–$8^{th}$</td>
<td>-.06</td>
<td>-.40 .26</td>
</tr>
<tr>
<td>Sahin (2013)$a$</td>
<td>230</td>
<td>$12^{th}$</td>
<td>.39</td>
<td>.05 .72</td>
</tr>
<tr>
<td>Sahin (2013)$b$</td>
<td>149</td>
<td>$12^{th}$</td>
<td>1.09</td>
<td>.72 1.46</td>
</tr>
<tr>
<td>Schmidt (2014)$a$</td>
<td>25</td>
<td>7th</td>
<td>.10</td>
<td>-.45 .66</td>
</tr>
<tr>
<td>Schmidt (2014)$b$</td>
<td>24</td>
<td>7th</td>
<td>.12</td>
<td>-.44 .70</td>
</tr>
<tr>
<td>Stoeger et al. (2013)</td>
<td>208</td>
<td>$12^{th}$</td>
<td>.06</td>
<td>-.13 .25</td>
</tr>
<tr>
<td>Tyler-Wood et al. (2012)</td>
<td>32</td>
<td>$4^{th}$–$5^{th}$</td>
<td>1.6</td>
<td>.72 2.50</td>
</tr>
<tr>
<td>Wyss et al. (2012)</td>
<td>72</td>
<td>6th–8th</td>
<td>.03</td>
<td>-.28 .35</td>
</tr>
</tbody>
</table>

Table 3 presents the mean effect sizes for each level of the different moderators, including OST program type, grade level, race, and gender. In table 3, when the 95% confidence interval does not include zero, the effect of the moderator is significantly different from zero. We also included the $Q_h$ values for the homogeneity analysis of the effect sizes for each moderator. A $Q_h$ value that is statistically significant indicates that the moderator influences the variation among the effect sizes. As indicated in Table 3, the effect sizes for OST program type (summer camp or after school) were both statistically significantly greater than zero.

However, based on the $Q_h$ statistic, OST program type was not a statistically significant moderator of student interest in STEM. For the analysis of student grade level, three studies were excluded due to overlapping grade levels (Heaverlo, 2011; Hughes et al., 2013; Innes et al., 2012). The effect sizes from grades 6–8, and 9–12 were statistically significantly different from zero; however, the effect sizes for K–5 were not. The $Q_h$ value for grade level was statistically significant, which indicates that grade level accounts for some of the variation in effect sizes. OST studies were coded as either academically focused or a combination of academic and socially oriented.
Table 3. Analysis of effect size moderators

<table>
<thead>
<tr>
<th>Moderator</th>
<th>$k$</th>
<th>$Q_B$</th>
<th>Effect Size</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time frame</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summer Camp</td>
<td>5</td>
<td>.29</td>
<td>[.06, .53]</td>
<td></td>
</tr>
<tr>
<td>After School</td>
<td>14</td>
<td>.39</td>
<td>[.18, .60]</td>
<td></td>
</tr>
<tr>
<td><strong>Grade Level</strong></td>
<td></td>
<td>23.61*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K–5</td>
<td>2</td>
<td>.77</td>
<td>[-.74, 2.28]</td>
<td></td>
</tr>
<tr>
<td>6–8</td>
<td>8</td>
<td>.16</td>
<td>[.01, .32]</td>
<td></td>
</tr>
<tr>
<td>9–12</td>
<td>5</td>
<td>.71</td>
<td>[.55, .87]</td>
<td></td>
</tr>
<tr>
<td><strong>Focus</strong></td>
<td></td>
<td>4.01*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic</td>
<td>7</td>
<td>.17</td>
<td>[.01, .34]</td>
<td></td>
</tr>
<tr>
<td>Social/Academic</td>
<td>12</td>
<td>.45</td>
<td>[.24, .65]</td>
<td></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td>.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female Only</td>
<td>7</td>
<td>.29</td>
<td>[.1, .47]</td>
<td></td>
</tr>
<tr>
<td>Male and Female</td>
<td>12</td>
<td>.36</td>
<td>[.15, .58]</td>
<td></td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td>.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unrepresentative</td>
<td>11</td>
<td>.39</td>
<td>[.17, .61]</td>
<td></td>
</tr>
<tr>
<td>Representative</td>
<td>8</td>
<td>.32</td>
<td>[.06, .59]</td>
<td></td>
</tr>
<tr>
<td><strong>Design</strong></td>
<td></td>
<td>16.52*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>4</td>
<td>.68</td>
<td>[.50, .85]</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>6</td>
<td>.17</td>
<td>[-.01, .34]</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>9</td>
<td>.36</td>
<td>[.15, .56]</td>
<td></td>
</tr>
</tbody>
</table>

Note: $k$ represents the number of effect sizes, *represents a statistically significant value of $Q_B$

The value of the $Q_B$ statistic for program focus was statistically significant, thus program focus accounts for some of the variability observed. Furthermore, both effect sizes for focus were statistically significantly greater than zero, and a larger effect size was observed for studies with a dual focus (academic and social). Although much of the literature supports the notion that interest in STEM is differentiated by race and gender, neither factor yielded a statistically significant value for the $Q_B$ statistic. Finally, the $Q_B$ value for the design quality (High, Medium, or Low) was statistically significant, thus the quality of the research design also accounts for some of the variability present in the analysis. Additionally, the largest effect sizes were observed for studies coded as high quality opposed to the medium, and low quality studies.

Limitations

During the evaluation and analysis process of this synthesis, several limitations were identified. First the quality of the studies presented in this synthesis varied from high to low, with the majority of the studies identified as lacking in the area of design quality or data representation. These ratings were based on several design features, as well as the presentation of program details necessary to assess the treatment fidelity. For instance, studies that utilized randomization of students into control and comparison groups received higher ratings than quasi-experimental studies. However, it is important to recognize that oftentimes researchers must submit to the will of the participants, which may prohibit the implementation of specific design protocols. Finally, given the nature of OST research it is impossible to completely control student activities after school; thus there is no such thing as a control or “no treatment group” (Miller, 2003). This issue was also addressed in many of the studies included in the synthesis as a limitation of their study. Subsequently, these limitations do not negate the findings presented in the current study, but provide further context for the interpretation of the evidence and implications provided.

Conclusion

The results of this synthesis lead to several conclusions and implications for praxis related to OST programs and their effect on student interest in STEM (Hall, Dickerson, Batts, Kauffmann, & Bosse, 2011; House, 2000; Turner, Steward, & Lapan, 2004). First, based on the results of this synthesis we conclude that OST can have a positive effect on the student interest in STEM. These results are analogous to the findings of similar synthesis with a focus on student achievement (Cooper et al, 2000; Lauer et al., 2006). Prior research recognized the effects of OST as a positive contributing factor to the academic success of students in mathematics, while our study provides an assessment of the extended benefit of OST programs on student interest in STEM. This
distinction is important because are results also suggest that STEM programs that are exclusively academic are less effective in promoting student interest in STEM. In relation to our first research question we conclude that OST programs have a small to medium positive effect on student interest in STEM based on common effect size benchmarks (Cohen, 1988). Thus, as posited by Bell et al. (2009), supplementary OST programs may help students consider STEM majors and subsequent careers. The results of this study quantify the affect of OST on student interest in STEM, but the instructional and experiential benefits are numerous for students that engage in STEM OST programs.

Barron et al. (2012) suggest that students in OST programs gain access to resources such as: (a) objects, (b) instruments, (c) expertise, and (d) settings not otherwise afforded to them. OST STEM programs can provide hands-on learning activities, situated in a realistic context, that reflect the technological advances to which students are accustomed outside of the traditional education setting (AAAS, 2001; Gallant, 2010). Unfortunately, student access to STEM OST programs in many parts of the country remains intangible (Adams et al., 2008). The lack of access to these programs prevents many students from opportunities to gain interest in STEM learning. Thus, policymakers should seek to provide equitable access to STEM programs across the nation. Access should be addressed both geographically and socioeconomically because many urban areas provide STEM enrichment programs, but the fees may exclude many traditionally underserved populations. The results of this study also suggest that the variation in positive effects of OST on student interest in STEM were influenced by several factors.

The timeframe of the programs was not a significant moderator of the program effects; however, the focus of the programs was a significant moderator worth further contextualization. The focus of each study included in this synthesis was identified as either academic or academic and social. The results suggest that OST programs with an academic and social focus had a larger effect on student interest in STEM, based on effect size magnitude. In general OST programs are designed with the flexibility to intertwine emotional social elements into the learning activities in a manner that is not easily accommodated by schools (Bevan & Michalchik, 2013). Yet, pressure to provide measures of academic effectiveness for funding agencies may cause some researchers and program coordinators to focus on achievement rather than more holistic measures of program effectiveness. Given the results of the present study it is appropriate that researchers, educators, and parents compare their personal goals to the goals of STEM OST programs to ensure that the program will sufficiently meet their needs.

The results from this study suggest that STEM interest is not sufficiently developed in OST settings that lack a social focus. As educators, researchers, and policymakers seek to design OST programs to develop STEM talent, a more complete list of skills should be considered. Dierking (2007) suggests that in addition to skills in STEM content, programs should develop leadership skills and interactions between the OST settings, student homes, communities, and the business sector. These activities can foster a social and emotional connection that students can draw from as motivation to pursue and complete degrees in STEM fields. Furthermore, women and students of color do not recognize STEM fields as platforms to reach their altruistic goal of helping others, which contributes to their decision not to choose a STEM-related career (Bonous-Hammarch, 2000; University of the Sciences, 2012). Women and students of color represent two sets of study characteristics that previous research consistently recognizes as contributing factors to STEM interest.

Despite substantial research to support race and gender as possible moderators of the effects of OST programs on STEM interest (Anderson & Kim, 2006; Carlone & Johnson, 2007; Durlak et al., 2010; Tan et al., 2013; Young, Young, Hamilton, 2013), the results of this study suggest that the variation in effect sizes was not statistically significantly influenced by either factor. One explanation for the results is that currently most studies refrain from disaggregating race and gender into independent subgroups (Larke et al., 2011). For example, all students of color are typically aggregated into the same group and groups are coded as homogeneous versus heterogeneous in meta-analysis research (Hembree & Dessart, 1986; Li & Ma, 2010). This is problematic because African American and Asian students are typically placed in the same group (heterogeneous), and thus the effect of interventions on individual racial groups is masked. Unfortunately, this problem exists because few studies disaggregate groups by race and gender. Thus, meta-analysts are constrained to examining racial and gender characteristics through the silhouette of homogenous and heterogeneous groups. In order to better inform future studies and examine the influences of race and gender in research synthesis, the disaggregation of results by race and gender is necessary (Capraro et al., 2009). The final program characteristic examined as a moderator was student grade level.

Consistent with previous research, grade level was a statistically significant moderator of the effects of OST programs on student interest in STEM. Although educators, researchers, and policymakers suggest that early access to STEM is important for the development of a strong STEM foundation (Duschl et al., 2007), the results
of this study suggest that OST program effects require further investigation at earlier grade levels. Grade level effect sizes were not statistically significantly different from zero for K–5 studies, but were statistically different from zero for the other grade spans, which is consistent with research that suggests the adolescent years are crucial for STEM interest development and maintenance (Caleon & Subramaniam, 2008; Frome et al., 2006). Middle school and high school represent two unique points of transition for the development of STEM learners. Middle school represents a transition in terms of preparation for learners and begins a decline in female student achievement, confidence, and interest in STEM (Else-Quest et al., 2010; George et al., 1992). Appropriately, middle school represents the most appropriate stage to initiate interventions designed to develop lasting interest in STEM that can be sustained through high school (Christensen et al., 2014). Based on the results of this study, we propose that more research is necessary to assess the effects of OST programs on STEM interest in the early grades, but more resources should be invested to develop and sustain middle and high school OST STEM programs.

The final significant moderator of OST program effects on STEM interest was study quality. This is a substantial finding because study fidelity is a major component of the validity and reliability of studies, and given the nature of OST programs fidelity can be difficult to achieve. Based on the results of this synthesis, study quality accounts for some of the variation in the effects of OST programs on student interest in STEM. High quality studies had a large effect size that was statistically significantly different from zero, while a small to medium size effect was observed for studies identified as low in quality. The medium quality studies were not statistically significantly different from zero. The results of this synthesis echo the call for more rigorous studies to assess the effects of OST on student interest in STEM (Adams et al., 2008; Lee, 2008). Specifically, the results of this synthesis suggest that well-designed studies are more effective means of promoting STEM interest through OST programs.

**Recommendations**

In order to remain competitive, the U.S. must develop a robust STEM workforce (Carenvale et al., 2011). Nonetheless, a STEM literate nation is the primary goal of the STEM agenda because STEM learning fosters 21st century skills that are essential for success in STEM and non-STEM fields. The general view of the world is that STEM education prepares students for the jobs of the future (Asunda, 2011). OST programs in STEM provide learning experiences that most traditional educational settings lack the means to emulate. The results of this study suggest OST programs have a positive effect on student interest in STEM. Based on these results, future studies should seek to identify specific academic and social elements of OST programs that may affect student interest in STEM. In addition, more work is necessary to better understand how OST programs can support the development of STEM interest across diverse populations. This information would help educators, researchers, and policymakers design STEM OST programs to meet our nation’s educational, social, and economic needs.

**References**


**Author Information**

**Jamaal Young**  
University of North Texas  
1155 Union Circle #310740, Denton, TX 76203  
Contact e-mail: jamaal.young@unt.edu

**Nickolaus Ortiz**  
Texas A&M University  
4232 TAMU College Station, Texas USA 77843

**Jemimah Young**  
University of North Texas  
1155 Union Circle #310740, Denton, TX 76203