ISSN 1303-0485 • **eISSN** 2148-7561
DOI 10.12738/estp.2016.1.0072

Educational Sciences: Theory & Practice • 2015 December • 15(6) • 1517-1529

Received | May 4, 2015
Accepted | June 18, 2015
OnlineFirst | September 3, 2015

Examining the Role of Inclusive STEM Schools in the College and Career Readiness of Students in the United States: A Multi-Group Analysis on the Outcome of Student Achievement

Niyazi Erdogan[a]
Balıkesir University

Carol Stuessy[b]
Texas A&M University

Abstract

The most prominent option for finding a solution to the shortage of workers with STEM knowledge has been identified as specialized STEM schools by policymakers in the United States. The current perception of specialized STEM schools can be described as a unique environment that includes advanced curriculum, expert teachers, and opportunities for internships and immersion. This study highlights the college readiness of STEM school graduates in comparison with traditional high school graduates. Using 11th grade students' high-stake test results in reading, mathematics, and science, this article compares the achievement outcomes of both school types. In answering the research questions related to student success for attendees of either STEM or traditional schools, this research concluded that success with reading, mathematics, and science high-stake tests for students does not differ by school type. However, student demographic variables (gender, ethnicity, socioeconomic status, and special education status) may influence the success of students attending STEM schools. For example, the results revealed a statistical significance between the reading, mathematics, and science scores of male, Hispanic, White, and economically disadvantaged students from STEM and traditional schools.

Keywords: STEM education • Specialized STEM schools • Achievement outcomes • College readiness

a **Correspondence author**
Niyazi Erdogan (PhD), Department of Elementary Education, Faculty of Education, Balikesir University, NEF Dinkciler Mah. Soma Cad. Merkez Balikesir 10100 Turkey
Research areas: STEM education; Project based learning; Engineering education; Teacher education; STEM schools
Email: niyazierdogan84@gmail.com & niyazierdogan@balikesir.edu.tr

b Assoc. Prof. Carol Stuessy (PhD), Department of Teaching, Learning and Culture, Texas A&M University, College Station, TX 77843 US
Email: c-stuessy@tamu.edu

Occupations in the 21st century increasingly require science, technology, engineering, and mathematics (STEM) knowledge (National Research Council [NRC], 2011). This demand is projected to continue during the next decade. However, the education system in the U.S. has not prepared enough students to fill those occupations requiring STEM knowledge (National Science Foundation [NSF], 2012; U.S. Department of Commerce, 2011). Until recently, U.S. businesses have managed to fill these occupations by importing students from other countries. However, this strategy has become outdated because of increased opportunities for similar occupations in other countries (Atkinson, Hugo, Lundgren, Shapiro, & Thomas, 2007). As a result, the shortage of workers with STEM knowledge has caused stress for U.S. businesses.

Policymakers in the U.S., realizing the importance of the situation, have developed new strategies for increasing the number of students to fill occupations requiring STEM knowledge (NRC, 2011). The first of these strategies includes: (*a*) improving the degree of training for STEM related careers, (*b*) increasing the number of people for the workforce, and (*c*) generating a more scientifically literate population (NRC, 2011). With these and other strategies, specific recommendations for increasing students include: (*a*) the creation of state-level mathematics and science standards, (*b*) recruitment and training of 100,000 STEM teachers over the next decade, (*c*) recognition for STEM teachers, (*d*) expansion of educational technology, (*e*) creation of extra-curricular opportunities for students, (*f*) creation of 1,000 new STEM-focused schools, and (*g*) provision of strong and strategic leadership (President's Council of Advisors on Science and Technology [PCAST], 2010). The PCAST authors identified specialized STEM schools as the most prominent recommendation.

In the last decade, most stakeholders (i.e., education leaders, policymakers, and researchers) have agreed that specialized STEM schools provide an optimum way for addressing the issue of reform for STEM education within the U.S. education system (Erdogan & Stuessy, 2015). In describing these schools, the NRC adopted a typology for identifying specialized schools. The NRC (2011) categorized specialized STEM schools under three headings: (*a*) selective STEM schools, (*b*) inclusive STEM schools, and (*c*) schools with STEM-focused career and technical education (CTE). Selective STEM schools serve students with aptitude and interest in STEM knowledge. These schools have certain admission criteria (e.g., past academic achievement; NRC, 2011; Subotnik, Tai, & Almarode, 2011). Inclusive STEM schools serve similar students; however, these schools have no admission criteria (NRC, 2011; Young et al., 2011). Schools with STEM-focused CTE serve students that are at risk for dropping out of school and accept students based on no criteria (NRC, 2011; Stone III, 2011).

Based on the above discussion, two problems arise to guide this study. The first problem is the blurred success of these schools at preparing students for college and career in STEM fields. Although a large amount of money has been invested in these schools, the success of these schools in preparing students is an unanswered question. The second problem involves better understanding of how student demographics correspond with student success on different achievement measures. These two problems suggest stakeholders have more to learn about the success of specialized STEM schools and the influence of student demographics over student performance on achievement measures.

The purpose of this study is to measure the college readiness of inclusive STEM high school (ISHS) graduates in comparison to traditional high school graduates. Schools classified as ISHS were chosen to represent a new school typology having the potential to direct females, minorities, and students with disabilities into STEM related careers. While evaluations and research are limited (Means, House, Young, Wang, & Lynch, 2013; Thomas & Williams, 2010; Young et al., 2011), policymakers continue to promote and expand STEM schools across the U.S. (PCAST, 2010). In order to explore the success of these schools, this study will be guided by the following research questions: (*1*) How do students from ISHS and traditional high schools in Texas compare regarding achievement outcomes in reading, mathematics, and science, and (*2*) for students attending ISHS in Texas, how do gender, ethnicity, socioeconomic status, and special education status associate with their achievement measures? Are these associations comparable to students attending traditional high schools in Texas?

Literature Review

The primary objective of all specialized STEM schools is to prepare students for college and careers in STEM fields, especially those students from historically underrepresented populations. To understand how these schools perform, Young et al. (2011) compared the achievement outcomes of students attending either inclusive STEM schools or traditional schools in Texas. When comparing students in 9th grade attending inclusive STEM or traditional schools, Young et al. found students from inclusive STEM schools performed slightly better on the mathematics high-stake test, were 1.8 times more likely to meet benchmarks for reading and mathematics high-stake tests, and were 0.8 times less likely to be absent from school. In addition, students in 10th grade attending inclusive STEM schools performed better on both mathematics and science high-stake tests and were 1.5 times more likely to meet the benchmarks for reading, mathematics, science, and history high-stake tests. Effect sizes indicated these differences, although statistically significant, were small. Finally, there were no statistically significant differences at any other grade level, suggesting limited benefit from inclusive STEM schools.

In another study, Means et al. (2013) compared students attending either inclusive STEM schools or traditional schools on their interest in STEM subjects and college matriculation. Results indicated students in 9th grade attending inclusive STEM schools were more interested in STEM subjects than similar students attending traditional schools. In addition, students attending inclusive STEM schools exhibited more confidence about earning high school and college diplomas than students attending traditional schools. Other findings in this study revealed students from inclusive STEM schools enrolled in more college preparatory courses within STEM disciplines, exhibited more interest in graduate school, and were more likely to enroll as STEM majors in college.

Thomas and Williams (2010), in another study, tracked students that had graduated from specialized STEM schools situated in the U.S. Of the 1,032 students in their study, 75% planned to continue formal education after high school and 40% planned to earn a doctorate degree. For these same participants, 51% pursued a science major and 10% pursued a mathematics major in college. Finally, findings showed 60% anticipated earning a STEM degree as college freshman while 55% had earned a STEM degree as college seniors.

T-STEM Schools

Currently, high schools in Texas serve over one million students of which at least 80% are categorized as Hispanic or White (Texas Education Agency [TEA], 2014b). The focus of this study is on the inclusive STEM schools initiative in Texas, which emphasizes the STEM education of historically underrepresented student populations. These schools also emphasize the students' college readiness and preparation for careers in STEM occupations. As a result of the STEM schools initiative, seven T-STEM schools were founded in Texas during the 2006-07 academic year. As of the 2013-14 academic year, 65 of these schools exist in Texas to serve a population of over 35,000 students. Funding for these schools has reached $133 million to date, which is more than the same size traditional schools receive, turning these schools into the largest investment for inclusive STEM high schools (ISHS) in the larger U.S. education system. Also, T-STEM schools are supported by partnerships with seven T-STEM centers, helping to create instructional materials and provide professional development workshops for over 2,800 teachers (TEA, 2013). The T-STEM schools were designed and implemented using a detailed blueprint, requiring students to (*a*) participate in a college preparatory curriculum, (*b*) develop real world relevant practices, (*c*) learn in a strong academic support system, and (*d*) master a wide range of STEM coursework (Avery, Chambliss, Pruiett, & Stotts, 2010; Corlu, 2013; Corlu, Capraro, & Capraro, 2014; NRC, 2011; Young et al., 2011). The primary objective in the mission statement for these schools is to prepare students for college and careers in STEM fields.

STEM Education within T-STEM Context

STEM education has been defined by many researchers and institutions (Merrill, 2009; Sanders, 2009; U.S. Department of Education, 2007). However, an agreement on the definition of STEM education has not been reached yet (Brown, 2012). Thus far, Merrill (2009) has suggested the only encapsulating definition based on the applications in the literature. This definition states STEM education is:

> "A standards-based, meta-discipline residing at the school level where all teachers, especially science, technology, engineering, and mathematics (STEM) teachers, teach an integrated approach to teaching and learning, where discipline-specific content is not divided, but addressed and treated as one dynamic, fluid study."

The key concept in this definition and many other definitions is the integrated approach to teaching and learning. T-STEM schools also emphasize integration of subjects in their mission statements. Although the blueprint for T-STEM schools does not offer a written definition of STEM education, one can define STEM education within T-STEM schools based on the blueprint as "an innovative approach that promotes integration of STEM subjects using inquiry, data collection, analyses, testing, technology usage, and problem solving to raise STEM-literate citizens and innovators," (Avery et al., 2010). In this definition, the blueprint defines a STEM-literate citizen as "one [who] understands how STEM can impact the quality of life for an individual, the education community, workforce of the future, the research environment, and public policy actions" (Avery et al., 2010, p. 40).

The integration of subjects in STEM education requires reform-based instructional strategies such as inquiry-based, project-based, and problem-based learning. Each instructional strategy aims to lead students to think critically, to innovate and invent solutions for problems that are faced daily (Avery et al., 2010). In these strategies, students have the chance to collaborate and apply what they have designed in real-world environments. Students are also provided with the opportunity to present their work to peers, teachers, and the community at large. For these instructional strategies to reach their aims, STEM faculty are expected to incorporate integrative content practices and research-based actions. However, this expectation comes with a number of disadvantages. First, one teacher has to know all. For example, a science teacher has to teach the science content along with the relevant technology, engineering, and mathematics. Also, teachers are expected to manage the process without misleading students because reform-based instructional strategies create a chaotic environment. The open-ended nature of these strategies can easily confuse students if teachers do not intervene timely. Finally, teachers are expected to use a blended form of formative and summative assessments for accurate instructional decisions in addressing the gaps in learning (Avery et al., 2010).

Reform-based instructional strategies incorporate multiple applications of STEM education in the T-STEM context. One of the well-known applications is robotics activity. Robots are great tools for structuring interdisciplinary instructions because they represent technology and engineering in one physical form and require mathematical knowledge at varying levels. In addition, robots can be used to run science experiments, thus completing the last part of the integrated STEM education (Erdogan, Corlu, & Capraro, 2013). Mega-structures are other well-known applications used in T-STEM schools. Mega-structures such as bridges and towers require physics and mathematical knowledge at different levels along with engineering skills (Ressler & Ressler, 2004). Technology in such instruction has the minor role but is still crucial for searching various building techniques. The other applications of STEM education in the T-STEM context have turned into national and international competitions such as a solar car challenge, water rocket challenge, catapult contest, gravity car race, and science Olympiad. Most STEM teachers in T-STEM schools encourage their students to participate in such competitions with projects they have designed in the classroom (Young et al., 2011). Through such competitions, STEM teachers aim to help students in developing the skills necessary for college and careers in STEM fields.

College and Career Readiness Standards

In 2010, the U.S. Department of Education set a clear goal for America's educational system, college, and career-ready high school graduates. However, state standards for college and career readiness did not align with the knowledge and skills necessary for post-graduation success. Statistics showed 40% of college freshman from both two and four-year institutions enroll in remedial courses (U.S. Department of Education, 2010). Although states also designed new assessments along with standards, these assessments were deemed inadequate at measuring students' knowledge and skills (U.S. Department of Education, 2010). To tackle the problem, the U.S. federal government developed a new approach. This approach included (*a*) supporting state standards for college and career readiness, (*b*) rewarding schools making progress, and (*c*) paying specific attention to the lowest-performing schools. The governments' first action in support of this approach was to reauthorize the Elementary and Secondary Education Act (ESEA; U.S. Department of Education, 2010). Essential changes in the ESEA included (*a*) rigorous standards in English language, arts, and mathematics, (*b*) reformed assessments aligned with college and career-readiness standards (CCRS), and (*c*) a structured reward system for schools and districts. Other changes in the ESEA recommended a support system, including: (*a*) improved support for teachers through professional development workshops, (*b*) enriched instruction for the lowest-performing schools, and (*c*) increased flexibility for schools and districts. The final recommendation for states was to continue implementing science standards and assessments (U.S. Department of Education, 2010). The efficacy of the new approach has yet to be evaluated in terms of preparing students for college and career.

In 2008, Texas focused on increasing the number of high school graduates who were college and career ready. Despite the progress Texan students have made in elementary and middle schools, the state trails other states in preparing students for college and career. Therefore, the Texas legislature passed the Advancement of College Readiness in Curriculum Bill (Educational Policy Improvement Center [EPIC], 2009). This bill required authorities to gather a team of experienced educators and university faculty to develop CCRS in English language, arts, mathematics, science, and social studies. The main objective of these standards was to help students gain the knowledge and skills necessary for college and career. Specifically, the courses designed with CCRS are intended to give students a set of core knowledge and skills across four subject areas (i.e., English, Social Studies, Mathematics, and Science; Sahin, Erdogan, Morgan, Capraro, & Capraro, 2012). According to the CCRS team, the more standards that students actualize, the more likely they are to be ready for college and career (EPIC, 2009).

Method

A quasi-experimental design was used to compare student outcomes from two different school types, T-STEM and traditional high schools (Campbell, Stanley, & Gage, 1963). In an attempt to answer the research questions listed above, achievement data was obtained through the Public Information Request system of TEA for 28,159 students in 11th grade attending one of 106 schools identified as either T-STEM or traditional. In addition, student demographic information was collected from the TEA using the same procedure. Student achievement was measured using scores from the Texas Assessment of Knowledge and Skills (TAKS) for reading, mathematics, and science. To examine associations between students' achievement and demographic variables, both descriptive and multi-group analyses were used.

Participants

The participants for analyses came from two separate data streams within the TEA. Participants included 28,159 students in 11th grade and 106 schools identified as either T-STEM or traditional. Student-level data included students' TAKS scores and demographic information. The TAKS scores represent standardized measures for students' mastery of reading, mathematics, history, and science content. Student demographic information included values for (*a*) gender, (*b*) ethnicity, (*c*) socioeconomic status, (*d*) English language proficiency, (*e*) English as a second language, (*f*) special education status, and (*g*) at-risk status. After obtaining the data for both participant sets, all data was compiled into two linked datasets.

Although 65 T-STEM schools have been founded under the Texas High School Project (THSP), only 53 such schools were identified from the student dataset. Data for students in the 11th grade from the student dataset were chosen because students in this grade take three of the four state achievement tests (i.e., reading, mathematics, and science). Variables of no concern for this study were removed from the dataset. Variables of concern were categorically coded, including students' gender (Female = 1, Male = 0), socioeconomic status (free meals, reduced-price meals, other economical disadvantages = 1, Not disadvantaged = 0), and special education status (Special education = 1, Not special education = 0). The variable for student ethnicity was dummy coded by declaring White ethnicity as the reference.

In a quasi-experimental study, results for the treatment group often find more meaning when a comparison of these results is conducted using data from a common or well-known group (Creswell, 2013). In our analyses, a sample of traditional schools from the entire population of Texas schools not designated as T-STEM but likely designated as high schools were used for comparative purposes. As a result, all schools (N = 8,529) in Texas were identified from the TEA website. After elimination of elementary, middle, charter, and alternative schools (i.e., night schools, T-STEM schools, early college high schools, recovery schools, and magnet schools), 1,309 schools remained. For analysis purposes, a sample of 53 schools was chosen from the population of 1,309 traditional schools serving students in 11th grade. A probability-stratified sampling procedure was applied by dividing T-STEM schools into four groups according to White student percentage (1 = 0-24%, 2 = 25%-49%, 3 = 50%-74%, 4 = 75%-100%). The first group had 34 T-STEM schools, the second group had nine such schools, the third group had seven schools, and the fourth group had three schools. The 1,309 traditional schools were grouped using the same method. From each group a similar number of traditional schools were randomly selected. After 53 traditional schools were identified, achievement and demographic data for students in the 11th grade from these schools were pulled from the TEA student dataset. Variables for comparison schools were also coded using the same method as in the T-STEM school dataset. In these two datasets, the new variable "STEM" was created to distinguish T-STEM schools from traditional schools. Finally, the two datasets were combined into one database for conducting analyses.

Tables 1 through 4 present cross distributions for students' school type and demographic categorizations (i.e., gender, ethnicity, socioeconomic status [SES], and special education status). Each of the tables provides information describing the relationships between students' school type and common demographic categorizations found in many education policy studies (Bozeman, Scogin, & Stuessy, 2013). The information in these tables suggests traditional schools serve more students but with similar distributions across the categorizations for gender, ethnicity, SES, and special education status.

Table 1
Cross Distribution of School Type and Gender

	Student Gender		
School Type	Male	Female	Total
Traditional	9,646	9,509	19,155
T-STEM	4,647	4,357	9,004
Total	14,293	13,866	28,159

Table 3
Cross Distribution of School Type and Socioeconomic Status

	Student Socioeconomic Status		
School Type	No	Yes	Total
Traditional	7,564	11,591	19,155
T-STEM	3,308	5,696	9,004
Total	10,872	17,287	28,159

Table 4
Cross Distribution of School Type and Special Education Status

	Student Special Education Status		
School Type	No	Yes	Total
Traditional	17,565	1,590	19,155
T-STEM	8,329	675	9,004
Total	25,894	2,265	28,159

Measurements

High-stake tests have been used for a number of decades to direct education policy (Heubert & Hauser, 1998). Results from high-stake tests have specifically been used to determine the success of students' schools, programs, and classrooms. These tests have also been used as indicators for students' college and career readiness. Until recently, the high-stake test accepted by most stakeholders in Texas was the Texas Assessment of Knowledge and Skills (TAKS; TEA, 2014a). TAKS measures students' achievement from 3rd grade through graduation across four academic disciplines (reading, mathematics, science, and social studies). However, these disciplines are not assessed at each grade. For example, students' achievement in the science discipline is assessed in 5th, 8th, 10th, and 11th grades. In this study, 11th grade was chosen because student achievement in three disciplines (reading, mathematics, and science) is assessed contemporaneously. The TAKS results for these three disciplines can be a useful indicator for making decisions regarding students' college and career readiness. The TEA replaced TAKS with the State of Texas Assessments of Academic Readiness (STAAR) in 2012. However, this replacement was progressive. As a result, in the 2012-13 academic year, 11th graders were still taking the TAKS exam. Therefore, TAKS results were used instead of STAAR.

Missing Data

As in most quasi-experimental studies using large datasets, this study has missing data. The missing data in this study was found within both independent and dependent variables. To address the missing data within independent variables, the listwise deletion method was chosen for 32 cases of missing data describing gender, ethnicity, SES, and special education. To address the missing data within dependent variables, four options were considered: (*a*) listwise deletion, (*b*) mean replacement, (*c*) maximum likelihood, and (*d*) multiple imputation. Multiple imputation was chosen as this method provides unbiased parameter estimates while ad-

Table 2
Cross Distribution of School Type and Ethnicity

	Student Ethnicity						
School Type	Asian	African American	Hispanic	Native American	Pacific Islander	Two or More Ethnicities	White	Total
Traditional	685	2,900	11,608	66	21	265	3,610	19,155
T-STEM	501	1,413	5,251	31	11	121	1,676	9,004
Total	1,186	4,313	16,859	97	32	386	5,286	28,159

dressing missing data (Graham, Olchowski, & Gilreath, 2007). Mplus version 7 was used to implement this method.

Data Analyses

Two analysis methods were used to answer the two research questions in this study: descriptive and multi-group analyses. For this purpose, Mplus 7 was used for calculating the means and standard deviations (*SD*) as well as conducting the multi-group analysis. Descriptive analysis was chosen to describe the center and spread of continuous data and the frequency distribution of categorical data. Multi-group analysis was chosen for the following reasons: (*a*) similar outcome variables for participants from different groups, (*b*) individual differences that remove the possibility of similarly responding to outcome measures, and (*c*) STEM applications that change the conceptual frame of reference against which a group responds to outcome measures over time (Muthén, 2002). When comparing student groups, the Wald test was used because of the test's robustness with large sample sizes. In this analysis, gender, ethnicity, SES, and special education variables were identified as independent variables while reading, mathematics, and science TAKS scores were identified as dependent variables.

As introduced in Figure 1, various student groups were compared using demographic variables. When comparing these groups, the Wald test was used because of the test's robustness with large sample sizes. If it is assumed there are K independent populations and $Y_{1j}, Y_{2j}, \ldots, Y_{nij}$ samples drawn from j^{th} population, where $j = 1, 2, \ldots, K$, $M_j = E(Y_{ij})$ will be the j^{th} population mean, where $i = 1, 2, \ldots, n_j$. Then, the formulation of these comparisons can be written as:

$$H_0: M_1 = M_2 = \ldots = M_K \qquad (1)$$

Where the alternative is H_1: $M_j = M_{j'}$, and $j = j'$. In the next section, the limitations of the methods described in this study are discussed.

Limitations

As with all studies, this one has multiple limitations. In this section, four limitations are identified. The first limitation is the absence of longitudinal data. The absence of data measuring students' college and career readiness prevents one from conducting a longitudinal study. However, longitudinal data of this nature is not readily available.

The second limitation is sampling of traditional schools. To complete the comparative analyses, 53 traditional schools were randomly selected. As criteria for stratified sampling, first the size and percentage of White students were used for each school. However, a large percentage of T-STEM schools were small. Therefore, only the percentage of White students was used when selecting the 53 traditional schools.

The third limitation is the categorization and definitions of certain variables. For example, SES had four categories in the original data (free meals = 1, reduced-price meals = 2, other economic disadvantages = 9, not identified as economically disadvantaged = 0). However, in our analyses, 1, 2, and

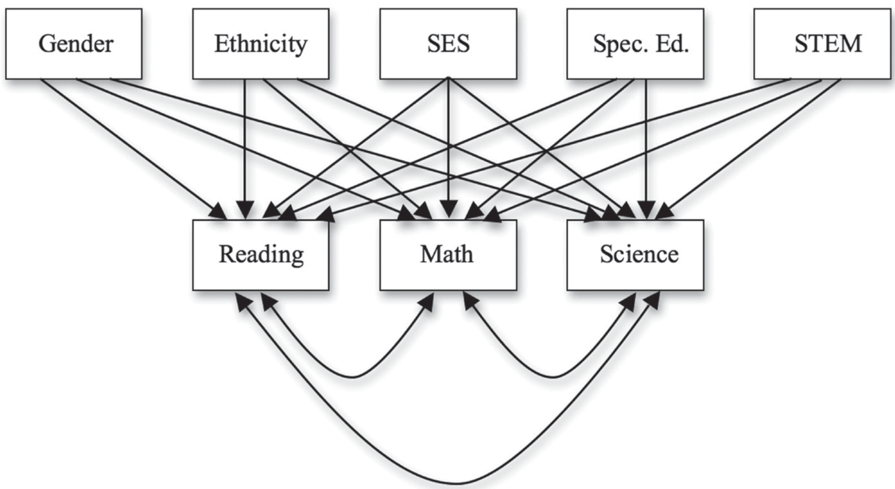

Figure 1. Diagram for multi-group model.

9 were categorized as economically disadvantaged and 0 as not disadvantaged to simplify discussion.

The fourth limitation is missing data. Approximately 25% of the data used in conducting this study was missing. To reduce bias and provide accurate results, a procedure was applied to the missing data. For reasons listed above, the multiple imputation method was chosen. Therefore, no data or information was lost regarding T-STEM and traditional schools' success at preparing students for college and career.

Although this study has limitations, it answered questions addressing the specialized STEM schools' success at preparing students for college and career using student level data. The data for 28,159 students in 11th grade from 53 T-STEM and 53 traditional schools included students' high-stake test (TAKS) scores and demographic information. To examine the association between variables, this study used descriptive and multi-group analyses. Missing data in this dataset was handled through the multiple imputation method. The next section presents the results of analyses.

Findings

Means, standard deviations, Wald statistics, and effect sizes were provided in Table 5 for the whole sample. Wald test results showed there was no statistically significant difference between students in traditional and T-STEM schools regarding reading, mathematics, and science scores ($Wald_{read} = .875$, $p = .350$; $Wald_{math} = 2.307$, $p = .129$; $Wald_{science} = .704$, $p = .402$). On average, students in T-STEM schools had higher scores for reading, mathematics and science ($mean_{read} = 2.555$, $mean_{math} = 2.253$, and $mean_{science} = 2.249$) than students in traditional schools ($mean_{read} = 2.245$, $mean_{math} = 2.228$, and $mean_{science} = 2.239$). However, these differences with the mean scores were not significant. Even though the Wald test results were not statistically significant, the effect sizes were calculated as reported in Table 5. For a detailed analysis, the sample was split into subgroups.

The means, standard deviations, Wald statistics, and effect sizes were provided in Table 6 for the gender subgroups. The Wald test results revealed a statistically significant difference between male (M) students in traditional and T-STEM schools regarding reading, mathematics, and science scores ($Wald_{maleread} = 6.132$, $p = .013$; $Wald_{malemath} = 10.295$, $p = .001$; $Wald_{malescience} = 7.058$, $p = .008$). For all three scores, male students in T-STEM schools performed better than male students in traditional schools. The results were similar for female (F) students except for science scores ($Wald_{femaleread} = 3.884$, $p = .049$; $Wald_{femalemath} = 6.619$, $p = .010$; $Wald_{female\text{-}science} = 3.424$, $p = .064$). The effect sizes were relatively small, ranging from 0.020 to 0.128. Each of the effect size values suggests higher mean scores on all TAKS tests for students in T-STEM schools when compared to students in traditional schools. The next analysis was run for ethnic subgroups.

The means, standard deviations, Wald statistics, and effect sizes were provided in Table 7 for ethnic subgroups. The Wald test results revealed statistically significant differences between Hispanic (H) and White (W) students in traditional and T-STEM schools for reading, mathematics, and science scores ($Wald_{H\_read} = 7.037$, $p = .008$; $Wald_{H\_math} = 11.743$, $p = .001$; $Wald_{H\_science} = 6.846$, $p = .009$; $Wald_{w\_read} = 5.217$, $p = .022$; $Wald_{w\_math} = 9.411$, $p = .002$; $Wald_{w\_science} = 6.250$, $p = .012$). However, the Wald test results were not significantly different among Asians (A), African Americans (AA), Native Americans (N), Pacific Islanders (P) and students from two or more (T) ethnic backgrounds in traditional and T-STEM schools. Although on average students in T-STEM schools performed better, these differences were not statistically significant. In fact, African American students and students from two or more ethnic backgrounds from traditional schools performed better than their counterparts in T-STEM schools regarding reading and science scores. The effect sizes for the Hispanic and White student subgroups were fairly small as well, ranging from .022 to .117. The next analysis was run for students' socioeconomic status.

Table 5
Cross Distribution of School Type and Average Score, Wald Statistic, and Effect Size on Achievement Measure for All Students

			Descriptive		Wald Statistic		
	N	School Type	*M*	*SD*	Score	*p*-value	Effect Size
Reading	19,155	Traditional	2,245	212	.875	.350	.047
	9,004	T-STEM	2,255	216			
Math	19,155	Traditional	2,228	236	2.307	.129	.104
	9,004	T-STEM	2,253	246			
Science	19,155	Traditional	2,239	204	.704	.402	.049
	9,004	T-STEM	2,249	208			

Table 6
Cross Distribution of School Type and Student Gender for Average Score, Wald Statistic, and Effect Size on Achievement Measure

				Descriptive		Wald Statistic		
	Gender	*N*	School Type	*M*	*SD*	Score	*p*-value	Effect Size
Reading	F	9,509	Traditional	2,263	209	3.884	.049	.042
	F	4,357	T-STEM	2,272	217			
	M	9,646	Traditional	2,226	214	6.132	.013	.060
	M	4,647	T-STEM	2,239	214			
Math	F	9,509	Traditional	2,228	226	6.619	.010	.081
	F	4,357	T-STEM	2,247	238			
	M	9,646	Traditional	2,228	245	10.295	.001	.128
	M	4,647	T-STEM	2,260	252			
Science	F	9,509	Traditional	2,233	194	3.424	.064	.020
	F	4,357	T-STEM	2,237	201			
	M	9,646	Traditional	2,246	214	7.058	.008	.065
	M	4,647	T-STEM	2,260	214			

Note. F represents female students and M represents male students.

Table 7
Cross Distribution of School Type and Student Ethnicity for Average Score, Wald Statistic, and Effect size on Achievement Measure

				Descriptive		Wald Statistic		
	Ethnic	*N*	School Type	*M*	*SD*	Score	*p*-value	Effect Size
Reading	A	685	Traditional	2,327	237	.012	.914	.025
	A	501	T-STEM	2,333	240			
Math	A	685	Traditional	2,372	272	1.765	.184	.192
	A	501	T-STEM	2,424	269			
Science	A	685	Traditional	2,334	246	.817	.366	.076
	A	501	T-STEM	2,352	228			
Reading	AA	2,900	Traditional	2,209	202	.403	.526	-.024
	AA	1,413	T-STEM	2,204	209			
Math	AA	2,900	Traditional	2,168	220	1.643	.120	.041
	AA	1,413	T-STEM	2,177	211			
Science	AA	2,900	Traditional	2,203	193	.935	.334	-.010
	AA	1,413	T-STEM	2,201	181			
Reading	H	11,608	Traditional	2,241	193	7.037	.008	.067
	H	5,251	T-STEM	2,254	196			
Math	H	11,608	Traditional	2,220	216	11.743	.001	.095
	H	5,251	T-STEM	2,241	224			
Science	H	11,608	Traditional	2,229	186	6.846	.009	.058
	H	5,251	T-STEM	2,240	193			
Reading	N	66	Traditional	1,531	197	.334	.563	.112
	N	31	T-STEM	1,553	195			
Math	N	66	Traditional	1,608	226	.348	.550	.174
	N	31	T-STEM	1,647	222			
Science	N	66	Traditional	1,696	200	.208	.648	.087
	N	31	T-STEM	1,713	191			
Reading	P	21	Traditional	1,815	215	.532	.466	.271
	P	11	T-STEM	1,871	197			
Math	P	21	Traditional	2,033	226	.390	.532	.236
	P	11	T-STEM	2,085	215			
Science	P	21	Traditional	2,039	188	.297	.586	.200
	P	11	T-STEM	2,077	192			
Reading	T	265	Traditional	2,194	263	.010	.922	-.097
	T	121	T-STEM	2,169	255			
Math	T	265	Traditional	2,201	286	.213	.644	.021
	T	121	T-STEM	2,195	280			
Science	T	265	Traditional	2,215	264	.000	.997	-.128
	T	121	T-STEM	2,183	236			
Reading	W	3,610	Traditional	2,291	234	5.217	.022	.043
	W	1,676	T-STEM	2,301	233			
Math	W	3,610	Traditional	2,292	256	9.411	.002	.117
	W	1,676	T-STEM	2,323	273			
Science	W	3,610	Traditional	2,300	225	6.250	.012	.022
	W	1,676	T-STEM	2,305	227			

Note. A represents Asian students, AA represents African American students, H represents Hispanic students, N represents Native American students, P represents Pacific Islander students, T represents students from two or more ethnic background, and W represents White students.

Table 8
Cross Distribution of School Type and Student Socioeconomic Status for Average Score, Wald Statistic, and Effect Size on Achievement Measure

				Descriptive		Wald Statistic		
	SES	*N*	School Type	*M*	*SD*	Score	*p*-value	Effect Size
Reading	N	7,564	Traditional	2,282	221	3.830	.050	.074
	N	3,308	T-STEM	2,298	221			
	Y	11,591	Traditional	2,221	203	6.141	.013	.044
	Y	5,696	T-STEM	2,230	210			
Math	N	7,564	Traditional	2,273	246	6.729	.001	.133
	N	3,308	T-STEM	2,307	264			
	Y	11,591	Traditional	2,199	224	11.286	.001	.105
	Y	5,696	T-STEM	2,222	229			
Science	N	7,564	Traditional	2,279	213	3.038	.081	.038
	N	3,308	T-STEM	2,287	222			
	Y	11,591	Traditional	2,214	194	8.271	.004	.067
	Y	5,696	T-STEM	2,227	196			

Note. Y represents students who are economically disadvantaged and N represents students who are not.

The means, standard deviations, Wald statistics, and effect sizes were provided in Table 8 for the socioeconomic subgroups. Wald test results showed statistically significant differences between economically disadvantaged (Y) students in traditional and T-STEM schools for reading, mathematics, and science scores ($Wald_{Y\_read}$ = 6.141, p = .013; $Wald_{Y\_math}$ = 11.286, p = .001; $Wald_{Y\_science}$ = 8.271, p = .004). On average, economically disadvantaged students in T-STEM schools performed better than their counterparts in traditional schools. However, this is not true for other (N) students except for the mathematics score ($Wald_{N\_read}$ = 3.830, p = .050; $Wald_{N\_math}$ = 6.729, p = .001; $Wald_{N\_science}$ = 3.038, p = .081). The effect sizes for economically disadvantaged students were again small, ranging from .044 to .105. The next analysis was run for students' special education status.

The means, standard deviations, Wald statistics, and effect sizes were provided in Table 9 for the special education subgroups. The Wald test results revealed no statistically significant difference between special education (Y) students in traditional and T-STEM schools regarding reading, mathematics, and science scores ($Wald_{Y\_read}$ = 2.550, p = .110; $Wald_{Y\_math}$ = 3.140, p = .076; $Wald_{Y\_science}$ = 1.400, p = .237). On average, special education students in T-STEM schools performed slightly better than their counterparts in traditional schools for reading and mathematics scores. However, special education students in traditional schools performed slightly better on science scores. The results were similar for other (N) students for reading scores but not for mathematics and science scores ($Wald_{N\_read}$ = 3.499, p = .061; $Wald_{N\_math}$ = 7.550, p = .006; $Wald_{N\_science}$ = 4.192, p = .041). The effect sizes for special education students were small, ranging from -.015 to .025.

Table 9
Cross Distribution of School Type and Student Special Education Status for Average Score, Wald Statistic, and Effect Size on Achievement Measure

				Descriptive		Wald Statistic		
	Spec. Educ.	*N*	School Type	*M*	*SD*	Score	*p*-value	Effect Size
Reading	N	17,565	Traditional	2,260	202	3.499	.061	.044
	N	8,329	T-STEM	2,269	208			
	Y	1,590	Traditional	2,075	250	2.550	.110	.020
	Y	675	T-STEM	2,080	242			
Math	N	17,565	Traditional	2,246	224	7.550	.006	.105
	N	8,329	T-STEM	2,270	234			
	Y	1,590	Traditional	2,039	278	3.140	.076	.025
	Y	675	T-STEM	2,046	284			
Science	N	17,565	Traditional	2,255	193	4.192	.041	.046
	N	8,329	T-STEM	2,264	196			
	Y	1,590	Traditional	2,076	251	1.400	.237	-.015
	Y	675	T-STEM	2,072	257			

Note. Y represents students who are in special education program and N represents students who are not.

The results reported in this section will be supported and explained with literature, then integrated into the theoretical framework in the conclusion.

Discussion

Stakeholders in STEM education recognize the need to address the issue of preparing students for college and career (Erdogan et al., 2013). However, the solution to this issue (i.e., specialized STEM schools) offered by stakeholders has yet to prove its value as a national resource. Hence, educational leaders across the nation are curious as to whether specialized STEM schools are outperforming traditional schools (Navruz, Erdogan, Bicer, Capraro, & Capraro, 2014). This study has presented a multi-group analysis to compare the students' achievement outcomes of traditional and T-STEM schools to understand the college and career readiness of students. A multi-group analysis provides opportunities to analyze similar outcome variables for different groups, distinguish individual differences, and take into account changes in the responses of a group over time because of an intervention (e.g., STEM curriculum; Muthén, 2002). In conducting the multi-group analysis results from the Wald test was preferred. The Wald test is a robust test when sample size is very large, as in this study.

Investments made in T-STEM schools can influence researchers' decision-making process. Many studies have suggested establishing new specialized STEM schools to address the issue of STEM education in the U.S. (Lynch, Behrend, Burton, & Means, 2013; Marshall, 2010; NRC; 2011; PCAST, 2010). Many researchers attempt to find differences between students in traditional and T-STEM schools because of these suggestions. Due to other null hypothesis tests having had inadequate controls for a large sample size, the Wald test was the best option for preventing Type I errors in such a comparison.

In response to research question one, based on data describing the state's high-stake test results, no statistically significant difference was found between students in traditional and T-STEM schools regarding reading, mathematics, and science scores. Although the Wald test results were not significant, the mean scores for students in T-STEM schools were higher than their counterparts in traditional schools. The effect-size values reflecting the mean differences were very small, however, confirming the earlier findings of Young et al. (2011).

One might expect reforms instituted in T-STEM schools would result in some significant differences in at least mathematics or science scores. There are a number of possible reasons that might have influenced the inability to find statistically significant differences across school types. For example, the model tested in the current study failed to account for differences in teachers across the school types. In addition, a large part of school effects may be related to student interest in STEM subjects rather than the overall improvement on students' scores regarding the high-stake tests (Means et al., 2013). Finally, one should consider that the T-STEM schools in this study were founded in urban areas and mostly populated by students from historically underrepresented populations, suggesting that these students' performance might be similar to students in traditional schools and actually could represent a benefit. This consideration requires further qualitative analysis.

In response to research question two, the student sample was broken into subgroups based on student demographics (e.g., gender). Results indicated that males in T-STEM schools, on the one hand, performed better than their counterparts in traditional schools regarding reading, mathematics, and science scores. On the other hand, females in T-STEM schools performed slightly better than their counterparts in traditional schools regarding reading and mathematics scores. However, the effect-size values for both subpopulations were very small. In addition, Hispanic and White students in T-STEM schools performed better than their counterparts in traditional schools with relatively small effect-size values. Other ethnic subpopulations did not exhibit any significant differences. Economically disadvantaged students in T-STEM schools also performed better than their counterparts in traditional schools. Once again, however, effect-size values were very small. Finally, students in the special education program from T-STEM schools showed no significant difference in regard to reading, mathematics, or science scores.

These achievement results between subgroups are promising because several target subpopulations (i.e., female, diverse, and disabled; NRC, 2011, NSF, 2013; PCAST, 2010; U.S. Department of Commerce, 2011) exhibited significant differences in achievement for specific subject areas. For example, female, Hispanic, and economically disadvantaged students' performance in comparison with their counterparts exhibited improvements regarding achievement scores in reading, mathematics, and science. However, work still remains for African Americans, Native Americans Pacific Islanders, students from two or more ethnic backgrounds, and students in the special education program.

Although work remains for these last student subgroups, one should consider that in Texas, Hispanic and White student populations constitute the majority of the total student population.

The college readiness of T-STEM graduates could be examined using various indicators. This study used results from reading, mathematics, and science high-stake tests because these indicators were readily available through state agencies. Despite the findings for T-STEM schools from this study, our knowledge of T-STEM schools' effects is limited. Other relevant variables may guide researchers to explore the successes of T-STEM schools at preparing students for college and career (Young et al., 2011). Also, cross-sectional research designs as employed in this study offer limited explanations of T-STEM schools' effects on the college readiness of students. Longitudinal research designs, however, offer a more powerful and stable explanation of these schools' effects (Willms & Raudenbush, 1989). Finally, the results from this study indicated an in-depth qualitative study of T-STEM schools' effect is required to understand the potential of these schools.

References

Atkinson, R. D., Hugo, J., Lundgren, D., Shapiro, M. J., & Thomas, J. (2007). Addressing the STEM challenge by expanding specialty math and science high schools. *NCSSSMST Journal, 12*(2), 14–23.

Avery, S., Chambliss, D., Pruiett, R., & Stotts, J. L. (2010). *T-STEM academy design blueprint, rubric, and glossary.* Retrieved from http://www.edtx.org/uploads/general/pdf-downloads/misc-PDFs/2011_TSTEMDesignBlueprint.pdf

Bozeman, T. D., Scogin, S., & Stuessy, C. (2013). Job satisfaction of high school science teachers: Prevalence and association with teacher retention. *Electronic Journal of Science Education, 17*(4), 1–19.

Brown, J. (2012). The current status of STEM education research. *Journal of STEM Education: Innovations and Research, 13*(5), 7–11.

Campbell, D. T., Stanley, J. C., & Gage, N. L. (1963). *Experimental and quasi-experimental designs for research.* Boston, MA: Houghton Mifflin.

Corlu, M. S. (2013). Insights into STEM education praxis: An assessment scheme for course syllabi. *Educational Sciences: Theory & Practice, 13*, 1–9.

Corlu, M. S., Capraro, R. M., & Capraro, M. M. (2014). Introducing STEM education: Implications for educating our teachers in the age of innovation. *Education and Science, 39*(171), 74–85.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches.* Thousand Oaks, CA: Sage.

Educational Policy Improvement Center. (2009). *Texas college and career readiness standards*. Austin, TX: University Printing Services at University of Texas-Austin.

Erdogan, N., & Stuessy, C. L., (2015). Modeling successful STEM high schools in United States: An ecology framework. *International Journal of Education in Mathematics, Science and Technology, 3*(1), 77–92.

Erdogan, N., Corlu, M. S., & Capraro, R. M. (2013). Defining innovation literacy: Do robotics programs help students develop innovation literacy skills? *International Online Journal of Educational Sciences, 5*(1), 1–9.

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science, 8*(3), 206–213.

Heubert, J. P., & Hauser, R. M. (Eds.). (1998). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academies Press.

Lynch, S. J., Behrend, T., Burton, E. P., & Means, B. (2013, April). *Inclusive STEM-focused high schools: STEM education policy and opportunity structures*. Paper presented at the annual conference of National Association for Research in Science Teaching (NARST), Rio Grande, Puerto Rico.

Marshall, S. P. (2010). Re-imagining specialized STEM academies: Igniting and nurturing decidedly different minds, by design. *Roeper Review, 32*(1), 48–60.

Means, B., House, A., Young, V., Wang, H., & Lynch, S. (2013). *Expanding access to STEM-focused education: What are the effects* [White paper]? Washington, DC: SRI International.

Merrill, C. (2009, February). *The future of TE masters degrees: STEM*. Paper presented at the 70th Annual International Technology Education Association Conference, Louisville, Kentucky.

Muthén, B. O. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, *29*(51), 81–118.

National Research Council. (2011). *Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics* (Committee on Highly Successful Science Programs for K-12 Science Education. Board on Science Education and Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education) Washington, DC: The National Academies Press.

National Science Foundation. (2012). *Science and engineering indicators 2012*. Arlington, VA: National Science Foundation (NSB 12-01).

National Science Foundation. (2013). *Women, minorities, and persons with disabilities in science and engineering: 2013* (Special Report NSF 13-304). Arlington, VA: National Science Foundation.

Navruz, B., Erdogan, N., Bicer, A., Capraro, R. M., & Capraro, M. M. (2014). Would a STEM school 'by any other name smell as sweet'? *International Journal of Contemporary Educational Research, 1*(2), 67–75.

President's Council of Advisors on Science and Technology. (2010). *Prepare and inspire: K-12 education in science, technology, engineering, and math (STEM) for America's future*. Washington, DC: Author.

Ressler, S. J., & Ressler, E. K. (2004). Using a nationwide Internet-based bridge design contest as a vehicle for engineering outreach. *Journal of Engineering Education, 93*(2), 117–128.

Sahin, A., Erdogan, N., Morgan, J., Capraro, M. M., & Capraro, R. M. (2012). The effects of high school course taking and SAT scores on college major selection. *Sakarya University Journal of Education, 2*(3), 96–109.

Sanders, M. (2009). STEM, STEM Education, STEMmania. *The Technology Teacher, 68*(4), 20–26.

Stone III, J. R. (2011, May). *Delivering STEM education through career and technical education schools and programs.* Paper prepared for the National Academies Board on Science Education and Board on Testing and Assessment for "Highly Successful STEM Schools or Programs for K-12 STEM Education: A Workshop", Washington D.C.

Subotnik, R. F., Tai, H. R., & Almarode, J. (2011, May). *Study of the impact of selective SMT high schools: Reflections on learners gifted and motivated in science and mathematics.* Paper prepared for the National Academies Board on Science Education and Board on Testing and Assessment for "Highly Successful STEM Schools or Programs for K-12 STEM Education: A Workshop," Washington D.C.

Texas Education Agency. (2013). *Texas science, technology, engineering, and mathematics initiative* (T-STEM). Retrieved from http://www.tea.state.tx.us/index2.aspx?id=4470&menu_id=814

Texas Education Agency. (2014a). *Texas assessment of knowledge and skills (TAKS) resources.* Retrieved from http://www.tea.state.tx.us/student.assessment/taks/

Texas Education Agency. (2014b). *Enrollment in Texas public schools 2012-13.* Retrieved fromhttp://www.tea.state.tx.us/acctres/Enroll_2012-13.pdf

Thomas, J., & Williams, C. (2010). The history of specialized STEM schools and the formation and role of the NCSSSMST. *Roeper Review, 32*(1), 17–24.

U.S. Department of Commerce. (2011). *STEM: Good jobs now and for the future* (ESA Issue Brief 03-11). Washington, DC: Economics and Statistics Administration.

U.S. Department of Education. (2007). *Report of the academic competitiveness council*. Washington, DC: Education Publications Center.

U.S. Department of Education. (2010). *A blueprint for reform: The reauthorization of the elementary and secondary education act*. Alexandria, VA: Education Publications Center.

Willms, J. D., & Raudenbush, S. W. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability. *Journal of Educational Measurement, 26*(3), 209–232.

Young, M. V., House, A., Wang, H., Singleton, C., SRI International, & Klopfenstein, K. (2011, May). *Inclusive STEM schools: Early promise in Texas and unanswered questions.* Paper prepared for the National Academies Board on Science Education and Board on Testing and Assessment for "Highly Successful STEM Schools or Programs for K-12 STEM Education: A Workshop", Washington D.C.