A Comparative Analysis of Secondary School STEM Research Programs in a Chinese School and an American School

Xiang Gong and Erik R. Mohlhenrich
Princeton International School of Mathematics and Science
xiang.gong, erik.mohlhenrich@prismsus.org

Abstract – **Countries around the world are committed to cultivating outstanding talent through STEM education. It is widely acknowledged that authentic STEM research programs are one of the most effective ways to achieve the goals of STEM education. In this paper, we present survey results in the 2018-2019 school year from school-based research programs at Princeton International School of Mathematics and Science (PRISMS) in the US and the High School Affiliated to Renmin University (commonly abbreviated as RDFZ) in China. A factorial MANOVA and a General Linear Model Univariate Analysis were used to test for similarities and differences between students' gains in dimensions of gains in thinking and working like a scientist (WIS), personal gains related to research work (PG), gains in skills (SKILL), attitudes or behaviors as a researcher (ATT), and career and graduate education aspirations (INF). Across both programs, we find significant gains on all variables as students' progress through their research experience. Scores from PRISMS students on WIS, PG, and ATT are significantly higher than those from RDFZ students. SKILL and INF showed significant correlations and thus were analyzed together; PRISMS students also scored higher on these variables. PRISMS 12th graders scored the highest of all school/grade level combinations. The results of this comparison speak to the efficacy of both programs in achieving the pedagogical goals of STEM research experiences. Variables that may have influenced the difference in outcomes between PRISMS and RDFZ are discussed, with particular attention given to the differences in the student population and school in general, number of students per project, and length of the research experience.**

Index Terms – comparative analysis, learning gains, research programs, secondary school, STEM education

INTRODUCTION

In the 21st century, it is widely acknowledged that strong science, technology, engineering, and mathematics (STEM) education is essential to a modern economy, and data show the need for STEM knowledge and skills will only grow in the future [1]. Therefore, countries around the world attach great importance to STEM education. Education authorities and policymakers in various nations have formulated their unique action plans for STEM education at the national level. For example, as early as 3 decades ago, the National Science Board of the United States released the *Rising Above the Gathering Storm*, made recommendations in different areas including K-12 STEM education and research. In 2010, the *Rising Above the Storm Revisited* became a wake-up call, analyzing how the original recommendations have or have not been acted upon, and giving rise to a revolutionary vision in the National Research Council's *A Framework for K-12 Science Education* [2],[3].

Several studies have noted that authentic STEM research programs are one of the most effective ways to achieve the goals of STEM education [4]–[9]. In particular, recent years have seen increased attention being given to various models of STEM research programs at secondary school level such as apprenticeship models, STEM research summer camps, placing high school students in university research teams, and school-based STEM research programs [10]–[14]. Numerous studies demonstrate the efficacy of these research programs in achieving the pedagogical goals of STEM education.

In China, STEM education has always played a pivotal role in national education development plans, as China's basic education reform emphasizes quality education focusing on the cultivation of the innovative spirit and practical ability through STEM subjects. In 2001, the Ministry of Education of China launched a nationwide high school curriculum reform plan which included STEM research-based learning programs [15]. The reason for this implementation was the education authorities in China realized that it was difficult to implement independent research-based learning programs (such as a STEM capstone project) in the regular subject teaching process due to the limitations of the teaching objectives, content, time and teaching methods. Thus, considering the importance of the research-based learning program, it was necessary to incorporate such programs as a compulsory element in the National High School Curriculum Plan. Unlike the diversified models in the United States, the Chinese model was mainly based on school research programs [16]. In 2017, the Ministry of Education of China revised the high school curriculum outline again, which specifically emphasized students' core competencies and how students could obtain the core competences in their learning and researching process. In light of this action, a steadily growing amount of work is beginning to focus on students' responses to learning gains from the research programs [17].

Although there are many studies on the efficacy and outcomes of secondary school STEM research programs, there is a relative lack of cross-cultural studies on these programs. In this article, we present survey results from school-based research programs at Princeton International School of Mathematics and Science (PRISMS) in the US and the High School Affiliated to Renmin University (commonly abbreviated as RDFZ) in China. We analyze the similarities and differences between student responses from these programs and discuss what can be learned from the results.

PRISMS Research Program and RDFZ Research Program: an Overview

Princeton International School of Mathematics and Science is a small international boarding school (100 students currently). In recent years, the nationality of PRISMS students has been approximately 70% Chinese, 25% American, and the rest from a variety of Asian and European nationalities. Students' research experience at PRISMS starts with the BASE project (Bridging the Arts through Science and Engineering) completed during the freshman year. In this program, each PRISMS 9th grader will work in teams to develop and conduct a semester-long research project, such as an analysis of the effects of LED light versus sunlight on the growth of kale (*Brassica oleracea* L. var. *acephala*) [18]. This program gives students the opportunity to develop real-world problem solving and research skills, take part in direct cross-curricular collaborations, and learn how to properly document and present results. Then, during their sophomore year, students continue to learn research concepts in AP STEM courses and select the lab in which they will pursue their research over the following two years. The STEM research program at PRISMS begins in junior year and lasts until graduation. Research is treated just like any other class; juniors and seniors have a research period totaling 3.5 hours per week. At the end of their senior year, each student will have produced a detailed journal documenting their progress, presented posters at school events and science fairs, given multiple presentations about their project, and written a technical paper describing their research.

PRISMS students are expected to participate in all aspects of the inquiry process throughout their research project — generating research questions, experimental design, data analysis, and communicating results. Given the diversity of subjects in which students can pursue research (see Table 1), the day-to-day nature of the work varies considerably between students.

TABLE 1
Selected Topics of Students' Researches from 2018-2019

No.	Students' Research Topics and Labs
1	Revealing Band Formation Factors of Agarose Gel Electrophoresis in a Two Buffer System (Biology Lab)
2	The Improvement of Biomolecules-Coated Titanium Dioxide Nanoparticles as Sunscreen Materials (Chemistry Lab)
3	Demonstrating the Existence of Dark Matter Using Data of the Galaxy Rotation Curve (Physics Lab)
4	LiDAR Based Obstacle Avoidance (Engineering Lab)
5	Dust Fall Prediction Based on Data Analysis Using Neural Network (Computer Science Lab)
6	Mobile Sensor Networks: Bounds on Capacity and Complexity of Realizability (Mathematics Lab)
7	Analysis of Academic Delay of Gratification (Statistics Lab)
8	The Effect of Music Tempo on Working Memory and Attention Span (Psychology Lab)

The High School Affiliated to Renmin University of China (RDFZ) is a high school under the jurisdiction of the Ministry of Education of China and one of the first batch of Beijing municipal model schools. RDFZ has been offering the research-based learning program for 26 years. Back in 1994, RDFZ formally launched a "Scientific Practice Course" for the 10th grade. After more than two decades of development, the research-based learning program has formed a complete course system as part of the high school curriculum plan. Every year, the research-based learning program is offered to all 11th graders. Students involved in the programs are required to have one hour class per week, as well as two hours of research study outside of school. In the first semester of the 11th grade, each student is required to learn the basic norms and procedures for doing research. They need to complete progressive assignments including literature review, observational study report, questionnaire design, etc. Moreover, students are expected to form research groups of 3-6 participants according to their own interests. They complete their one-year research programs under the guidance of a mentor. Some of the mentors are full-time teachers specialized in research-based learning programs, and the others are subject teachers at RDFZ.

TABLE 2
Students' STEM-related Research Areas at Rdfz

No.	Research groups
1	Theory of machine learning and pattern recognition
2	Physics
3	Microbiology
4	Plant physiology and Molecular biology
5	Enzyme and molecular biology
6	Chemical analysis and preparation
7	Geography and environment
8	Energy System
9	Electronic information engineering
10	Computer science and information technology
11	Modeling and design
12	Virtual Reality Technology
13	Imaging and future media
14	Robotics
15	Science and Technology Research club
16	Maker space
17	Astronomy
18	Information and Communication Engineering

At the end of the junior year, there is a “Research Day” where each group of students is required to present a poster to demonstrate their research project. The students will also make presentations, and awards will be selected by university professors invited by the school and mentors responsible for research-based learning.

FIGURE 1
RDFZ RESEARCH POSTER SESSION

RESEARCH QUESTIONS

During the final exam week of the 2018-2019 school year, the URSSA (Undergraduate Research Self-Assessment Survey) was taken by all students at PRISMS and randomly selected students at RDFZ in order to evaluate their learning outcomes from the research programs. Five independent variables were assessed with the survey results: (1) gains in thinking and working like a scientist (WIS); (2) personal gains related to research work (PG); (3) gains in skills (SKILL); (4) attitudes or behaviors as a researcher (ATT), and (5) career and graduate education aspirations (INF). Our research questions were generated from these five variables.

(1) Were there significant interactions among the two schools (PRISMS versus RDFZ) and three grade levels (10^{th} grade through 12^{th} grade) with regard to the linear combination of students’ responses to WIS, PG, SKILL, ATT, and INF?

(2) Was there a statistically significant difference between the two schools with responses to the linear combination, or individuals of WIS, PG, SKILL, ATT, and INF?

(3) Were there statistically significant differences among responses from 10^{th} graders through 12^{th} graders with regard to the linear combination, or individuals of WIS, PG, SKILL, ATT, and INF?

METHODS

All 82 students at PRISMS (first-year 9^{th} graders through fourth-year 12^{th} graders) participated in the survey. The survey was administered at the end of the 2018-2019 school year after all research activities. After the survey, we found one freshman student sequentially chose 1-2-3-4-5 as he went through the questions and this student’s survey was excluded. Meanwhile, the high school students across 10^{th} grade through 12^{th} grade from RDFZ, who were participating in or have completed the research-based learning programs were randomly selected to conduct the URSSA survey. Since Chinese high schools were only in grades 10-12, freshman surveys at PRISMS were excluded from analysis in order to conduct parallel comparisons. The sample selection from RDFZ was based on the consideration of the factor of cost, time, and effectiveness. The final population of the research study was 330, including 54 10^{th} graders (34 males and 20 females), 169 11^{th} graders (80 males and 89 females), and 107 12^{th} graders (58 males and 49 females). All students were assigned an ID number for the survey and were anonymous to the authors of the paper.

TABLE 3
DISTRIBUTION OF THE RESEARCH STUDENTS

		School		Total
		PRISMS	RDFZ	
Grade	10th	24	30	54
	11th	20	149	169
	12th	17	90	107
Total		61	269	330

As previously mentioned, PRISMS and RDFZ conducted a pilot implementation of the URSSA online survey during the 2018-2019 school year. All questions from the online survey were adopted with permission from the URSSA team. The URSSA survey was modeled on the SALG (Student Assessment of Their Learning Gains) instrument, with an emphasis on student reports of their own learning gains in cognitive, behavioral, and affective areas [19]. Students are required to answer 134 questions grouped in 17 blocks from the survey. In terms of students’ learning outcomes, there are 8 dependent variables describing WIS, 6 dependent variables describing PG, 12 dependent variables describing SKILL, and 8 dependent variables describing ATT. Regarding INF, there were some questions that were not relevant for high school students and thus we excluded these from further analysis.

After the data from 330 participants were collected, the IBM SPSS 18 package was used to process the data. Two primary statistical methods in the package were used in this paper. First, a factorial MANOVA was designed in this study to determine whether there were significant interaction or significant differences among the two schools (PRISMS versus RDFZ) and three grade levels (10^{th} grade through 12^{th} grade), with regard to the linear combinations of WIS, PG, SKILL, ATT, and INF. Second, a GLM (General Linear Model) Univariate Analysis was also run to identify if there were a significant difference between the two schools and three grade levels, with regard to WIS, PG, SKILL, ATT, and INF, respectively.

RESULTS

The descriptive statistics for the five variables were shown in Table 4. All variables roughly followed the normal distribution. For instance, the P-P plots in Figure 2 indicated that the PG plot looked very linear, while the SKILL plot showed some slight curvature. Neither plot had strong outliers or clear skewness. Individual observations of each variable were independent.

TABLE 4
DESCRIPTIVE STATISTICS OF INDEPENDENT VARIABLES

	WIS	PG	SKILL	ATT	INF
Mean	3.9232	3.9347	3.8673	3.8129	3.4980
Std. Deviation	.94242	.95723	.99606	1.07252	1.25511

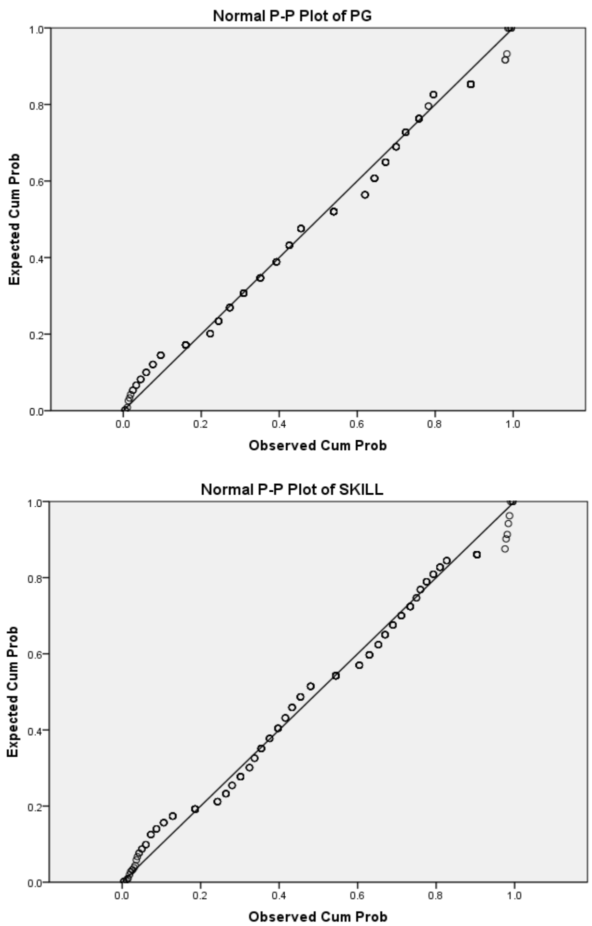

FIGURE 2
P-P PLOTS OF PG AND SKILL

Table 5 showed the Pearson correlation between SKILL and INF was 0.390, which revealed that there were significant interactions among the two schools and three grade levels with regard to the linear combination of students' responses to SKILL and INF. The factorial MANOVA in Table 6 stated that the Pillai's Trace=0.062, F(4, 648)=5.157, p=0.000, and multivariate eta^2=0.031. Meanwhile, there was a significant difference between PRISMS and RDFZ, with regard to the linear combination of students' responses to SKILL and INF, with Pillai's Trace=0.036, F(2, 323)=6.041, p=0.003, and multivariate eta^2=0.036. However, there was not a significant difference among the grade levels with regard to the linear combination of students' responses to SKILL and INF. Table 6 showed the indicators as follows: Pillai's Trace=0.028, F(4, 648)=2.314, p=0.056, multivariate eta^2=0.014.

TABLE 5
PEARSON CORRELATIONS AMONG FIVE VARIABLES

		WIS	PG	SKILL	ATT	INF
WIS	Pearson Correlation	1	.868**	.862**	.791**	.337**
PG	Pearson Correlation	.868**	1	.885**	.809**	.347**
SKILL	Pearson Correlation	.862**	.885**	1	.824**	.390**
ATT	Pearson Correlation	.791**	.809**	.824**	1	.349**
INF	Pearson Correlation	.337**	.357**	.390**	.359**	1

**. Correlation is significant at the 0.01 level (2-tailed).

TABLE 6
MULTIVARIATE TESTS

Effect		Value	Hypothesis df	Error df	Sig.	eta^2
Intercept	Pillai's Trace	.918	2.000	323.000	.000	.918
Grade	Pillai's Trace	.028	4.000	648.000	.056	.014
School	Pillai's Trace	.036	2.000	323.000	.003	.036
Grade * School	Pillai's Trace	.062	4.000	648.000	.000	.031

In order to analyze the significant interaction among the two schools and three grade levels, we created a new variable InterCombo including the six possible combinations from the two school groups and three grade levels, which were PRISMS & 10th grade, PRISMS & 11th grade, PRISMS & 12th grade, RDFZ & 10th grade, RDFZ & 11th grade, and RDFZ & 12th grade. Table 7 showed that there were significant differences among the six combinations for the linear combination of students' responses to SKILL and INF (p=0.000).

TABLE 7
MULTIVARIATE TESTS

Effect		Value	F	Hypothesis df	Sig.	eta^2
Intercept	Pillai's Trace	.918	1814.759	2.000	.000	.918
InterCombo	Pillai's Trace	.145	5.080	10.000	.000	.073

Table 8 showed the means and standard deviations of SKILL. The mean value of the PRISMS students from 10th grade to 12th grade increased from 3.873 to 4.726, whereas the mean of RDFZ students decreased from 3.969 to 3.610, then bounced back to 3.857. Numerically, the average of the PRISMS 10th graders was slightly lower than that of RDFZ 10th graders, but the 11th and 12th grades were both higher than RDFZ students. Turky HSD in Table 9 revealed that there were significant differences with regard to students' responses to SKILL between (1) PRISMS & 11th grade and RDFZ & 11th grade, with the p=0.008 and 95% Confidence Interval from 0.039 to 1.340; (2) PRISMS & 11th grade and RDFZ & 10th grade, with the p=0.008 and 95% Confidence Interval from 0.129 to 1.384; (3) PRISMS & 12th grade and RDFZ & 11th grade, with the p=0.000 and 95% Confidence Interval from 0.515 to 1.717; (4) PRISMS & 12th grade and RDFZ & 12th grade, with the p=0.012 and 95% Confidence Interval from 0.121 to 1.617. In general, among the combinations which were statistically significant, the PRISMS groups obtained higher scores than the RDFZ groups.

TABLE 8
DESCRIPTIVE STATISTICS OF SKILL

Six New Combinations	Mean	Std. Deviation
PRISMS &10th Grade	3.873	1.060
PRISMS &11th Grade	4.300	.752
PRISMS & 12th Grade	4.726	1.829
RDFZ &10th Grade	3.969	.745
RDFZ &11th Grade	3.610	.899
RDFZ &12th Grade	3.857	.812
Total	3.867	.996

TABLE 9
TUKEY HSD FOR SKILL

(I) Six New Combinations	(J) Six New Combinations	Sig.	95% Confidence Interval	
			Lower Bound	Upper Bound
PRISMS &11th Grade	RDFZ &11th Grade	0.030499	0.039214	1.340267
PRISMS & 12th Grade	RDFZ &10th Grade	0.0081245	0.128789	1.383909
PRISMS & 12th Grade	RDFZ &11th Grade	0.0000028	0.515098	1.716765
PRISMS & 12th Grade	RDFZ &12th Grade	0.0123251	0.120946	1.617149

In terms of INF, Table 10 showed the means and standard deviations. Grade 10 students in PRISMS scored slightly lower than those in Grade 10 in RDFZ, but their grades in Grades 11 and 12 surpassed those of RDFZ grades. Turky HSD in Table 11 indicated that there was a significant difference between PRISMS & 11th grade and RDFZ & 11th grade, with the p=0.004 and 95% Confidence Interval from 0.226 to 1.904.

TABLE 10
DESCRIPTIVE STATISTICS OF INF

Six New Combinations	Mean	Std. Deviation
PRISMS & 10th Grade	3.633	1.110
PRISMS &11th Grade	4.300	.638
PRISMS &12th Grade	3.700	1.334
RDFZ &10th Grade	3.902	1.249
RDFZ &11th Grade	3.234	1.277
RDFZ &12th Grade	3.416	1.182
Total	3.498	1.255

TABLE 11
TUKEY HSD FOR INF

(I) Six New Combinations	(J) Six New Combinations	Sig.	95% Confidence Interval	
			Lower Bound	Upper Bound
PRISMS &11th Grade	RDFZ &11th Grade	0.004265	0.226117	1.904084

As there were no significant interactions among the two school groups and three grade levels with regard to the linear combination of students' responses to WIS, PG, and ATT, a GLM Univariate Analysis was executed to determine whether there were significant interactions among the two school groups and three grade levels with regard to WIS, PG, and ATT, respectively. The descriptive statistics in Table 12 showed the means and standard deviations of WIS, PG, and ATT, respectively.

TABLE 12
DESCRIPTIVE STATISTICS OF WIS, PG, AND ATT

Grade	School	WIS		PG		ATT	
		Mean	S.D.	Mean	S.D.	Mean	S.D.
10th	PRISMS	4.117	.8866	3.973	1.027	3.786	.9790
	RDFZ	3.982	.7853	4.049	.724	3.994	.8055
11th	PRISMS	4.266	.6984	4.088	.738	3.856	.9676
	RDFZ	3.669	.7925	3.681	.854	3.549	.9072
12th	PRISMS	4.717	1.830	4.685	1.861	4.807	2.115
	RDFZ	4.029	.7664	4.122	.689	3.766	.9415
Total	PRISMS	4.402	1.3106	4.291	1.373	4.211	1.581
	RDFZ	3.814	.8009	3.853	.816	3.722	.898

With respect to WIS, further analysis showed that there was no significant interaction (p=0.214). However, there was a significant difference between the PRISMS and RDFZ (p=0.001). In other words, students from PRISMS and RDFZ displayed distinct differences in the nature of their reported gains in WIS. The scores of PRISMS students were significantly higher than the scores of the RDFZ students. Besides, a significant difference was found among 10th grade through 12th grade with p=0.039. A closer look at the means revealed that the WIS scores increased significantly from the 10th graders to 12th graders.

Regarding PG, the small p-value (p=0.159) suggested there was no significant interaction according to the GLM Univariate Analysis. However, there was a significant difference with respect to PG between PRISMS and RDFZ (p=0.031), which indicated the PG score of PRISMS students was significantly higher than the scores from the RDFZ students. Meanwhile, a significant difference (p=0.007) was also found among the three groups from 10th grade through 12th grade, which indicated that the PG scores increased significantly from the 10th graders to 12th graders.

With reference to ATT, the interaction effect was statistically significant (p=0.006). Moreover, there was a significant difference between PRISMS and RDFZ (p=0.014), which indicated the ATT score of PRISMS students was significantly higher than the scores from the RDFZ students. The difference among three grade levels (10th grade through 12th grade) was also found to be statistically significant with p=0.007. A follow-up Tukey HSD Test for ATT revealed that PRISMS & 12^{th} grade and PRISMS & 10^{th} grade, PRISMS & 11^{th} grade, RDFZ & 10^{th} grade, RDFZ & 11^{th} grade, RDFZ & 12^{th} grade all had significant differences, respectively. Furthermore, a significant difference was also found between RDFZ & 11^{th} grade and RDFZ & 12^{th} grade. The p-values and 95% Confidence Intervals were shown in Table 13.

TABLE 13
TUKEY HSD FOR ATT

(I) Six New Combinations	(J) Six New Combinations	Sig.	95% Confidence Interval	
			Lower Bound	Upper Bound
PRISMS & 10th Grade	PRISMS &12th Grade	0.023	-1.9533	-0.8708
PRISMS & 11th Grade	PRISMS &12th Grade	0.029	-1.8419	-0.0602
PRISMS & 12th Grade	RDFZ &10th Grade	0.008	0.1369	1.4888
PRISMS & 12th Grade	RDFZ &11th Grade	0.000	0.6106	1.905
PRISMS & 12th Grade	RDFZ &12th Grade	0.003	0.2348	1.8464
RDFZ &10th Grade	RDFZ &11th Grade	0.016	0.0521	0.8378

DISCUSSION

This study surveys two STEM research programs implemented by PRISMS in the United States and RDFZ in China. Students in the programs took the URSSA survey in the 2018-2019 school year and 5 variables were measured — (1) gains in thinking and working like a scientist (WIS); (2) personal gains related to research work (PG); (3) gains in skills (SKILL); (4) attitudes or behaviors as a researcher (ATT), and (5) career and graduate education aspirations (INF). Across both programs, we find significant gains on all variables as students' progress through their research experience. This result speaks to the efficacy of both programs in achieving the pedagogical goals of STEM research experiences.

There are a number of common threads running across the two programs that may explain the consistency of these results. First, unlike a large number of research programs that operate outside the formal STEM curriculum, the programs identified in the study are both fully integrated into the school's curriculum and is therefore intended for all students. Second, as with many high school research education initiatives, the programs identified in this study have similar content and educational approach. For instance, both programs offer students introductory research method courses or seminars to establish or advance their understanding of research through a critical exploration of research language, ethics, and approaches. When students are about to select research labs, they are required to draft proposals and meet with mentors to develop their research projects. After their entry to the labs, they need to write weekly journal entries, update progress with posters/presentations, prepare for research project exhibitions, and submit research papers at the culmination of the project.

TABLE 14
CHARACTERISTICS OF PRISMS AND RDFZ RESEARCH PROGRAMS

Characteristics	PRISMS	RDFZ
Central Program Theme	STEM-related	All subjects including STEM
Time Length in the Lab	2 years	1 year
Program Type	Mandatory to students	Mandatory to students
Freshmen Research Experience	Yes	No
Research Method Courses / Seminars Offered	Yes	Yes
Project-based Approach and Case-study Led Instruction	Yes	Yes
Number of Students per Project	1	3-6
Research Project Proposal	Yes	Yes
Weekly Journal Entries	Yes	Yes
Progress Poster/Presentations	Yes	Yes
Research Project Exhibition	Yes	Yes
Research Paper Submission	Yes	Yes
Mentors Source	Internal	Internal + External
Amount of third party involvement	Low	High
Inter-school partnership	No	Yes

In contrast, there are also notable differences between the results from the PRISMS and RDFZ programs (Table 14). Scores from PRISMS students on WIS, PG, and ATT are significantly higher than those from RDFZ students. SKILL and INF showed significant correlations and thus were analyzed together. In general, PRISMS students scored higher on these variables than their RDFZ counterparts; PRISMS 12^{th} graders scored the highest of all school/grade level combinations. Interpreting these differences requires caution as there are numerous variables that could influence the results which are not directly related to the research programs. Differences in the student population and school in general (PRISMS is much smaller and more selective) and/or cultural factors may explain the results. With that being said, a few key differences in the research programs could be partly responsible for the pedagogical outcomes. The PRISMS research program is an individual experience whereas students at RDFZ typically work in groups of 3-6 students. While PRISMS students certainly collaborate at times, they are expected to be self-directed and carry out all stages of research individually. The individual nature of the project may positively affect their perceived gains in these variables as students will feel a sense of ownership and autonomy over the direction of their project. Some of the difficulties that come with collaboration may harm RDFZ students' perceived gains; we can certainly imagine a proportion of students who feel as if they are not participating in all stages of the research process or perhaps the project doesn't truly match their interests. We can also imagine situations in which the collaborative nature of a project galvanizes a student who may otherwise not be particularly interested in research. Another key factor is the length of the research experience — 2 years at PRISMS and 1 year at RDFZ. Studies have shown the length of the research experience to be a significant factor in determining learning outcomes [4],[20],[21]. In previous work, we highlighted the advantages and disadvantages of the two-year research program at PRISMS [14].

Recent efforts at both schools have been focused on developing these research programs along three primary dimensions: (1) establishing and renovating the school-based STEM research laboratories, (2) enhancing cooperation with universities and research institutions, and (3) training and professional development of research mentors. The first element aims to provide greater capacity and allow for a deeper and wider range of student research projects. Progress on (2) gives students more opportunities to collaborate with university researchers and to participate in national and regional research competitions and conferences (including ISEC). The third dimension is dedicated to improving the ability of teachers to mentor research and achieve the pedagogical goals of the programs. Progress on this dimension includes providing professional development opportunities and supporting faculty to publish papers and give presentations at academic conferences. Looking forward, further progress on these dimensions should improve the research programs at both schools and sustain them for years to come.

This study also suggests several promising directions for future research. One limitation of this study is that the URSSA survey is aimed at undergraduate research experiences; many of the questions were not particularly relevant to students in our program. This points to a need for assessment models and relevant evaluation scales for high school research programs. Second, we find that both programs contain many of the same or similar components. Further research can verify the validity of these common characteristics through qualitative and quantitative methods, and analyze whether there is a correlation or cause and effect relationship between the common characteristics and the students' gains from research programs. Third, the differences between responses from PRISMS and RDFZ obtained in this research provide a rich basis for further research. Are the differences between the PRISMS and RDFZ results caused by cultural differences or features of the research program such as length and group versus individual projects? How does the integration of the research program into the school's curriculum affect the learning outcomes? Future research at PRISMS and RDFZ, along with large-scale work on STEM research programs across the United States and China, are needed to address these questions.

Acknowledgment

We wish to acknowledge all research students at PRISMS and RDFZ who participated in the URSSA survey in 2019. We appreciate the help of the administration and faculty of PRISMS and RDFZ and would like to especially thank Madam Pengzhi Liu, the Honorary Principal for RDFZ Union School and RDFZ for her invaluable support to this research.

References

[1] Courtney Tanenbaum, "STEM 2026: A Vision for Innovation in STEM Education." 2016. www.air.org/resource/stem-2026-vision-innovation-stem-education. Web. Accessed: January 28, 2020.

[2] National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. 2010. *Rising Above the Gathering Storm, Revisited*. Washington, DC: The National Academies Press.

[3] National Research Council. 2012. *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

[4] T. D. Sadler, S. Burgin, L. McKinney, *et al.* 2010. "Learning science through research apprenticeships: A critical review of the literature." *J. Res. Sci. Teach.*, vol. 47, no. 3, pp. 235–256.

[5] S. H. Russell, M. P. Hancock, and J. McCullough. 2007. "Benefits of undergraduate research experiences." *Science*. Vol. 316, Issue 5824, pp. 548-549.

[6] A. B. Hunter, S. L. Laursen, and E. Seymour. 2007. "Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development." *Sci. Educ.* Vol. 91. pp. 36-74.

[7] M. C. Linn, E. Palmer, A. Baranger, *et al.* 2015. "Undergraduate research experiences: Impacts and opportunities." *Science*. Vol. 347(6222), pp. 627.

[8] J. Bennett, L. Dunlop, K. J. Knox, *et al.* 2018. "Practical independent research projects in science: a synthesis and evaluation of the evidence of impact on high school students." *Int. J. Sci. Educ.*, vol. 40, no. 14, pp. 1755-1773.

[9] O. A. Adedokun, A. B. Bessenbacher, L. C. Parker, *et al*. 2013. "Research skills and STEM undergraduate research students' aspirations for research careers: Mediating effects of research self-efficacy." *J. Res. Sci. Teach.*, Vol. 50, no. 8, pp. 940–951.

[10] M. Musavi and C. James. 2015. "A Transformative Apprentice Research STEM Program (RTP – Strand 4)." *ASEE Annu. Conf. Expo. Conf. Proc.* Vol. 2015-June.

[11] A. Yousuf, M. A. Mustafa, M. M. Hayder, *et al*. 2018. "STEM education internship program." *ASEE Annu. Conf. Expo. Conf. Proc.* Vol. 2018-June.

[12] A. Sirinterlikci and S. F. Sirinterlikci. 2013. "Development of a Summer High School Research Program." *ASEE Annu. Conf. Expo. Conf. Proc.* Vol. 2013-June.

[13] Center of Excellence in Leadership for Learning at University of Indianapolis. 2009. "Summary of Research on Project-based Learning." *Cent. Excell. Leadersh. Learn.*, no. June, pp. 1–3.

[14] X. Gong and E. R. Mohlhenrich. 2019. "An Integrated Secondary School STEM Research Program: Results, Challenges, and Opportunities." *ISEC 2019 - Proc. 9th IEEE Integr. STEM Educ. Conf.*, Vol. 2019. pp. 76-82.

[15] Ministry of Education of China, "Notice of the Ministry of Education on Issuing the Guidelines for the Implementation of Research-based Learning in High Schools." 2001. www.moe.gov.cn/srcsite/A06/s3732/200104/t20010409_82009.html. Web. Accessed: January 28, 2020.

[16] H. Wang. 2002. "Open to self-construction: A research on research-based learning," Ph.D. dissertation, Faculty of Education, East China Normal University. Shanghai, 2002. Accessed on: December 28, 2019.

[17] Y. Li, A. H. Schoenfeld, A. Andrea, *et al*. 2019. "On Thinking and STEM Education." *Journal for STEM Education Research*. Vol. 2, Issue 1, pp. 1–13.

[18] E. R. Mohlhenrich, S. V. Samsonau, and R. P. Spencer. 2018.

"Integrating science through authentic research in secondary schools." *ISEC 2018 - Proc. 8th IEEE Integr. STEM Educ. Conf.*, vol. 2018-Janua, pp. 14–21.

[19] T. J. Weston and S. L. Laursen. 2015. "The undergraduate research student self-assessment (URSSA): Validation for use in program evaluation." *CBE Life Sci. Educ.* Vol. 14, no. 3, pp. 1–10.

[20] H. Thiry, T. J. Weston, S. L. Laursen, *et al*. 2012. "The benefits of multi-year research experiences: Differences in novice and experienced students' reported gains from undergraduate research." *CBE Life Sci. Educ*. 2012 Fall; Vol. 11(3), pp. 260-272.

[21] O. A. Adedokun, L. C. Parker, A. Childress, *et al.* 2014. "Effect of time on perceived gains from an undergraduate research program." *CBE Life Sci. Educ.* 2014 Spring; Vol. 13(1), pp. 139-148.

AUTHOR INFORMATION

Xiang Gong, Mathematics Teacher & Research Mentor, Princeton International School of Mathematics and Science
Erik R. Mohlhenrich, Biology Teacher & Research Mentor, Princeton International School of Mathematics and Science