APPLIED LEARNING STUDENT QUESTIONNAIRE: *OVERALL ANALYSIS*

Overall Results
May 2015

Executive Summary

Participants and Methods

In May 2015, 970 students across 6 Race to the Top programs completed the Applied Learning Student Questionnaire (ALSQ). The response rates displayed in Table 1 suggest that 78% of the total number of participating students responded to the survey. The response rates per program ranged from 12% (RT3 Computational Thinking) to 100% (Tift County Mechatronics). Although there is no agreed-upon standard for a minimum response rate, Martella, Nelson, Morgan, and Marchand-Martella (2013)[1] suggest that a response rate of 50% is *adequate* for analysis and reporting, 60% is *good*, and 75% or higher is considered *very good*. Overall, the response rate achieved across 6 Race to the Top programs is considered *very good* for reporting and analysis.

Table 1. Survey Response Rates

Program	# of Survey Respondents	Total # of Participating Students	Survey Response Rate
21st Century Rockdale County	301	341	88%
Real STEM Georgia Southern	215	334	64%
RT3 Computational Thinking	4	33	12%
STEM for Life Carroll County	239	325	74%
STEP Academy Gwinnett	140	144	97%
Tift County Mechatronics	71	71	100%
Total	**970**	**1,248**	**78%**

Note. The number of participating students represent approximations and may not reflect recent changes to the participant population (e.g., dropouts).

The ALSQ[2] is designed to measure pre and post gains related to student problem solving and communication skills, self-management and engagement. The ALSQ is a self-report questionnaire that includes 36 items to assess students' attitudes on the following survey constructs:

1. **Intrinsic Motivation**: motivation stemming from goals of mastery, learning and challenge. Example, "It is important for me to learn what is being taught in this program."
2. **Self-Management/Self-Regulation:** effortful and persistent behaviors that are used to guide, monitor, and direct the success of one's learning and performance. Example, "I turn all my assignments in on time."
3. **Intent to Persist:** aspirations, plans, and goals to pursue additional education and a career in STEM. Example, "I intend to get a college degree in STEM (Science, Technology, Engineering, and Math)."
4. **Problem Solving:** inquiry-based learning environment that provides higher-order cognitive tasks and real-world applications. Example, "I work out explanations on my own."

---

[1] Martella, R., Nelson, J., Morgan, R., & Marchand-Martella, N. (2013). *Understanding and Interpreting Education Research*. New York, NY: The Guilford Press.

[2] See Appendix A for information related to the construct reliabilities of the ALSQ.

Executive Summary, continued

5. **Implementation Activities:** hands-on activities designed to increase exposure to STEM topics and real-world applications. Example, "We learn what scientists/technicians/engineers/ mathematicians or other STEM professionals do."

Results & Discussion

- **ALSQ Survey Constructs**

  Table 2 summarizes students' responses to the ALSQ survey constructs across all programs. In aggregate, students show statistically significant increases in *Intrinsic Motivation*, *Self-Management/Self-Regulation* skills, and *Intent to Persist*. In addition to assessing statistical significance from "before" to "now," effect sizes—a measure of the magnitude of an intervention on students' attitudes—were computed. Specifically, effect sizes were computed using Cohen's *d* and are intended to measure the practical importance of a significant finding. Cohen (1988) classified effect sizes as small, *d*=0.2; medium, *d*=0.5; and large, *d*=0.8.[3] Table 2 suggests that medium effect sizes were found for *Intrinsic Motivation,* Self-*Management/Self-Regulation,* and *Intent to Persist*. Across all constructs, the largest effect size observed was for *Intrinsic Motivation* (*d*=0.59). This suggests that the programs were particularly effective at enhancing students' interests to learn and derive value from the material being taught. For example, prior to participating in the programs, only 57% of students said that understanding STEM is important to them compared to 80% after the program. See Table 4 for more information.

  To maximize impact, we would expect students' average scores to exceed 4.00 on a 5-point Likert scale (1, *Strongly Disagree* to 5, *Strongly Agree*). In light of this benchmark, it is important to note that the "now" scores across two constructs— *Intent to Persist* and *Implementation Activities*— did *not* reach or exceed the optimal average of 4.00. Figure 1 suggests that additional work may be needed in the above mentioned areas.

Table 2. Summary of Results by Constructs

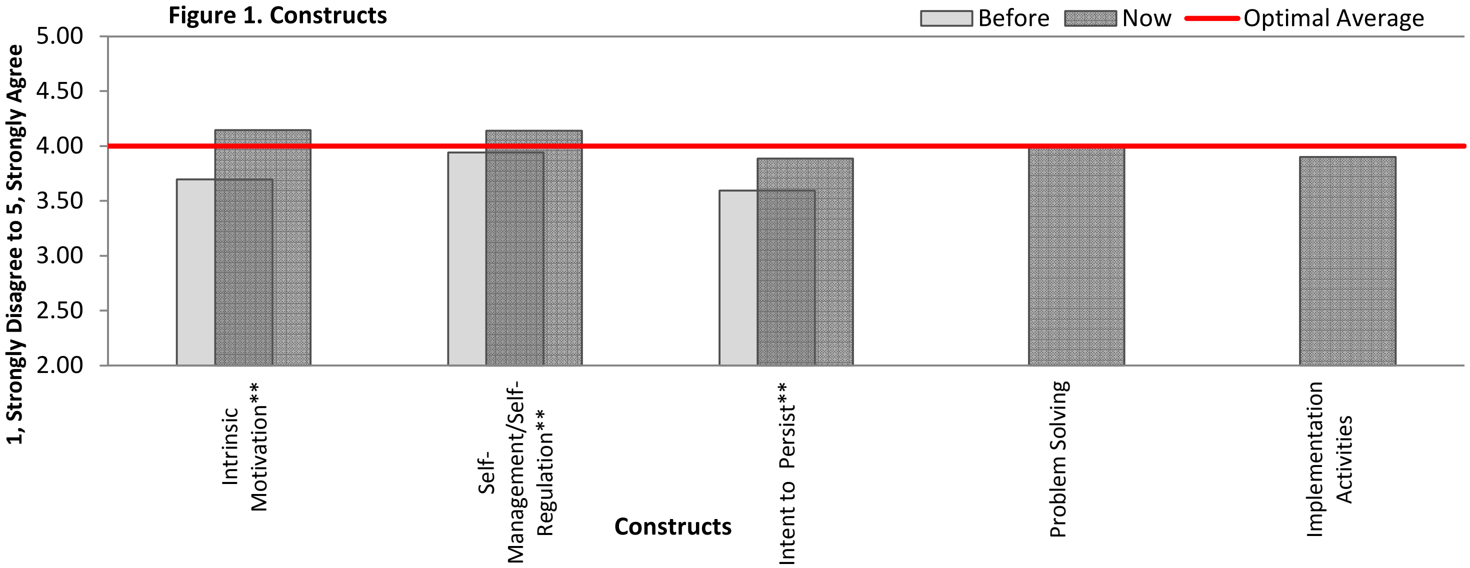
Overall- Constructs					
Constructs		**n**	**Mean**[1]	**Paired Samples t-test**[2]	**Effect Size (interpretation)**[3]
Intrinsic Motivation	Before	967	3.70	p<0.001**	0.59 M
	Now	960	4.15		
Self-Management/Self-Regulation	Before	966	3.94	p<0.001**	0.37M
	Now	962	4.14		
Intent to Persist	Before	964	3.59	p<0.001**	0.42 M
	Now	962	3.88		
Problem Solving	Now	961	4.00	N/A	N/A
Implementation Activities	Now	957	3.90	N/A	N/A

Note. Scale; 1, *Strongly Disagree* to 5, *Strongly Agree*. [1] Reference lines are set at 3.5 and 4. [2]Please note that only students with matched Pre and Post data were assessed for significance. Desired statistically significant changes are highlighted in green. **p<0.001, *p<0.01, †p<0.05. Negatively worded statements were reverse coded for mean computations. [3]Effect size (Cohen's *d*): Small (<.2); Medium (.2 to .8); Large (>.8). Small effect sizes are highlighted in light red; medium effect sizes are highlighted in dark orange; large effect sizes are highlighted in dark green.

[3] Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed). Hillsdale, NJ: Lawrence Earlbaum Associates.

Executive Summary, continued

Figure 1. Constructs



Note. **p<0.001, *p<0.01, †p<0.05; Scale is truncated for visual clarity.

- **ALSQ Survey Constructs by Program**[4]

Examining the ALSQ results by individual program, it is evident that across nearly all programs, students show statistically significant increases in *Intrinsic Motivation, Self-Management/Self-Regulation* and *Intent to Persist*. Examining effect sizes, students in the following two programs show medium to large effect sizes: Tift County Mechatronics and STEP Academy Gwinnett. This suggests that the above mentioned programs had a medium to large impact on students' attitudes. It is important to note that due to the very small sample size (n=4) for the RT3 Computational Thinking program, statistical power[5] was compromised and we were unable to compute t-tests and effect sizes.

Table 3. Summary of Results by Constructs per Program

Overall- Constructs per Program										
Constructs		**21st Century Rockdale County (n=301)**			**Real STEM Georgia Southern (n=215)**			**RT3 Computational Thinking (n=4)**		
		Mean	*t-test*	Effect Size	Mean	*t-test*	Effect Size	Mean	*t-test*	Effect Size
Intrinsic Motivation	Before	3.83	$p<0.001^{**}$	0.36^{M}	3.60	$p<0.001^{**}$	0.65^{M}	4.06	--	--
	Now	4.08			4.07			4.64		
Self-Management/ Self-Regulation	Before	4.04	$p<0.001^{**}$	0.22^{M}	4.09	$p<0.001^{**}$	0.44^{M}	4.39	--	--
	Now	4.14			4.24			4.32		
Intent to Persist	Before	3.72	$p<0.001^{**}$	0.22^{M}	3.56	$p<0.001^{**}$	0.46^{M}	3.85	--	--
	Now	3.85			3.83			4.25		
Problem Solving	Now	3.79	N/A	N/A	4.16	N/A	N/A	4.70	N/A	N/A
Implementation Activities	Now	3.66			3.92			4.45		

Note. Scale= 1, *Strongly Disagree* to 5, *Strongly Agree*. Desired statistically significant changes are highlighted in green and undesired statistically significant changes are highlighted in red. **p<0.001, *p<0.01, †p<0.05. Effect size (Cohen's d): Small (<.2); Medium (.2 to .8); Large (>.8). Small effect sizes are highlighted in light red; medium effect sizes are highlighted in dark orange; large effect sizes are highlighted in dark green.

[4] For additional information related to Real STEM Georgia Southern University see Appendix B.

[5] Statistical power is the ability of a test to detect an effect, if the effect actually exists. Statistical power is contingent on an adequate sample size and an adequate effect size (the salience of the treatment relative to the noise in measurement).

Executive Summary, continued

Continued, Table 3. Summary of Results by Constructs per Program

Continued, Overall- Constructs per Program										
Constructs		STEM for Life Carroll County (n=239)			STEP Academy Gwinnett (n=140)			Tift County Mechatronics (n=71)		
		Mean	*t-test*	Effect Size	Mean	*t-test*	Effect Size	Mean	*t-test*	Effect Size
Intrinsic Motivation	Before	3.68	$p<0.001^{**}$	0.53^{M}	3.51	p<0.001**	0.93^{L}	3.82	$p<0.001^{**}$	1.11^{L}
	Now	4.10			4.17			4.71		
Self-Management/ Self-Regulation	Before	3.82	$p<0.001^{**}$	0.35^{M}	3.66	p<0.001**	0.54^{M}	4.02	$p<0.001^{**}$	0.61^{M}
	Now	4.04			4.01			4.42		
Intent to Persist	Before	3.43	$p<0.001^{**}$	0.38^{M}	3.45	p<0.001**	0.64^{M}	3.98	$p<0.001^{**}$	0.84^{L}
	Now	3.71			3.92			4.70		
Problem Solving	Now	4.04	N/A	N/A	3.80	N/A	N/A	4.66	N/A	N/A
Implementation Activities	Now	3.97			3.80			4.74		

Note. Scale= 1, *Strongly Disagree* to 5, *Strongly Agree*. Desired statistically significant changes are highlighted in green and undesired statistically significant changes are highlighted in red. **p<0.001, *p<0.01, ⴕp<0.05. Effect size (Cohen's d): Small (<.2); Medium (.2 to .8); Large (>.8). Small effect sizes are highlighted in light red; medium effect sizes are highlighted in dark orange; large effect sizes are highlighted in dark green.

In order for programs to maximize their effectiveness, we would expect "now" scores to reach or exceed the optimal average of 4.00 on a 5-point Likert scale (1, *Strongly Disagree* to 5, *Strongly Agree*). Figures 2 – 6 display "now" scores for each program and construct. For example, Figure 2 indicates that 6 out of 6 programs met or exceeded the optimal average for *Intrinsic Motivation*. In general, programs not reaching or exceeding the red horizontal line may need additional attention. For instance, 4 out of 6 programs did not reach the optimal average for *Intent to Persist* and *Implementation Activities*.

Figure 2. Intrinsic Motivation ("Now" Scores)

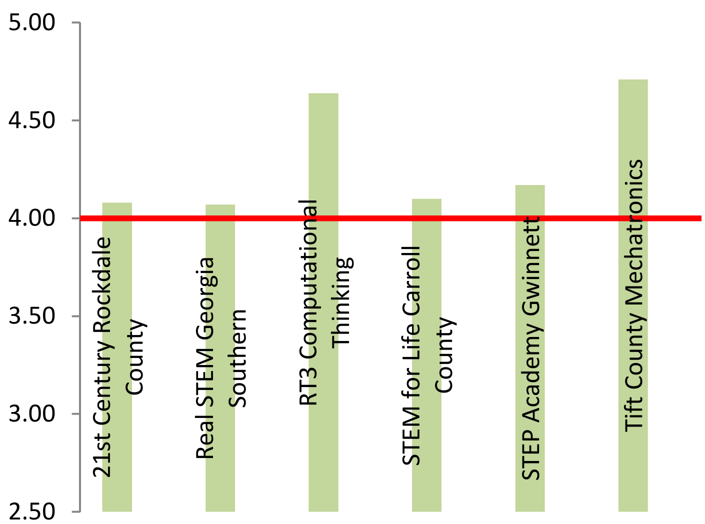


Figure 3. Self-Management/Self-Regulation ("Now" Scores)

Scale= 1, Strongly Disagree to 5, Strongly Agree. Scale was truncated for visual clarity. Programs that met or exceeded the optimal average of 4.00 are reflected in green; programs that fell below the optimal average are reflected in red.

Executive Summary, continued

Figure 4. Intent to Persist ("Now" Scores)

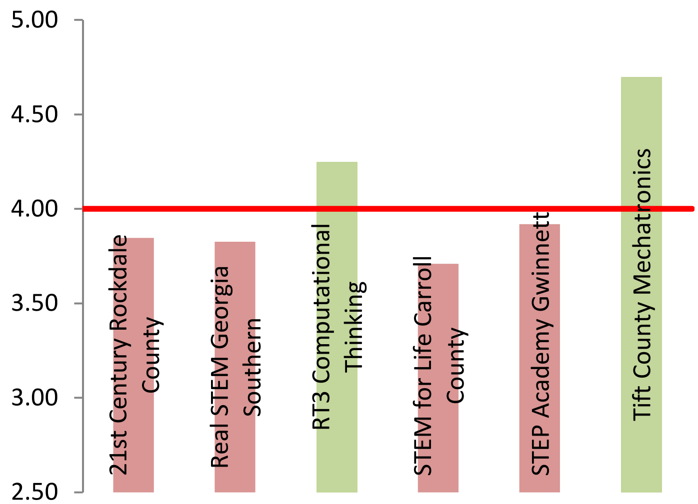


Figure 5. Problem Solving ("Now" Scores)

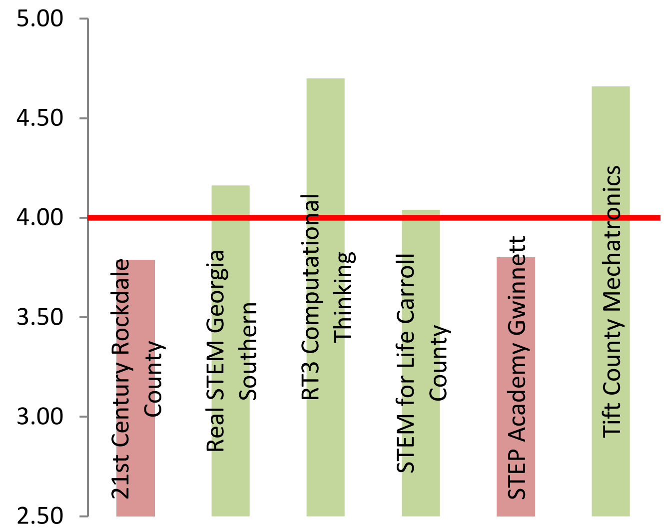


Figure 6. Implementation Activities ("Now" Scores)

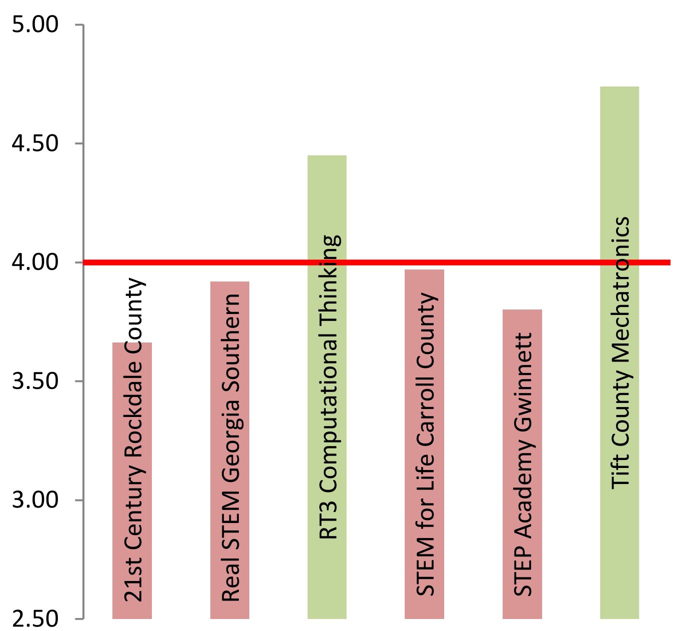
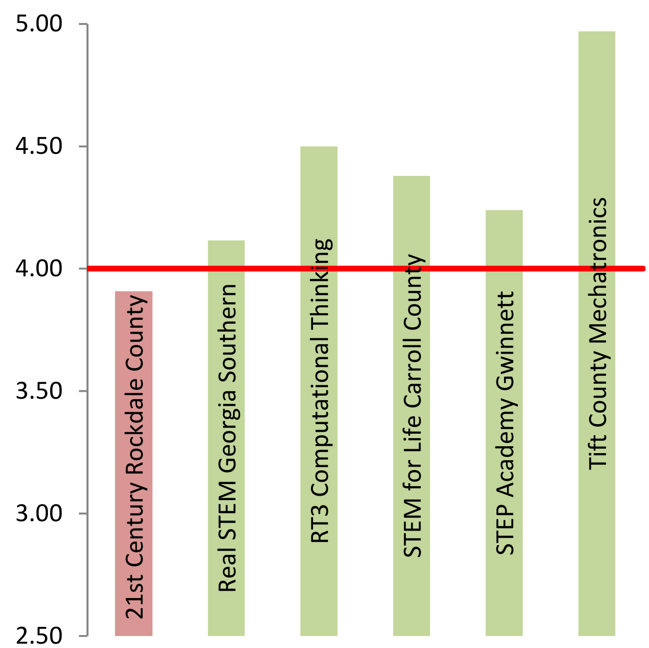


Figure 7. Overall Program Ratings



Scale= 1, *Strongly Disagree* to 5, *Strongly Agree*. Scale was truncated for visual clarity. Programs that met or exceeded the optimal average of 4.00 are reflected in green; programs that fell below the optimal average are reflected in red.

Executive Summary, continued

- **Program Rating**
  Collapsing across all programs, students' ratings exceeded the optimal average of 4.00. On a 5-point Likert scale where 1 signifies *Very Poor* and 5 signifies *Excellent*, the average score was a 4.20. See Table 12. Looking at Figure 7, it is evident that 5 out of 6 programs were rated above the optimal average. The 21st Century program in Rockdale County may need additional support to reach the optimal average.

- **Areas for Further Improvement**
  Across all programs, further enhancing implementation activities and students' intentions to persist in STEM may be warranted. Specifically, students' ratings suggest that the inquiry-based learning environment may be improved by allowing students more opportunity to choose their own topics, work out explanations on their own, interact with STEM professionals, and engage in topics with real-world applications. Such implementation activities may strengthen students' intentions and motivations to pursue additional education in STEM fields. Additionally, improving the response rates on the surveys for a couple of programs (e.g., RT3 Computational Thinking) may enhance the generalizability of the results. In particular, achieving at least a 50% response rate for each program may be necessary to improve the validity and reliability of the findings.

Table 4. Intrinsic Motivation

Intrinsic Motivation		n	Mean[1]	Paired Samples t-test[2]	1 (Strongly Disagree)	2 (Disagree)	3 (Neutral)	4 (Agree)	5 (Strongly Agree)
1. I prefer class work that is challenging so I can learn new things.	Before	966	3.47	p<0.001**	5%	12%	31%	33%	18%
	Now	959	3.94		4%	4%	20%	39%	33%
2. It is important to me to learn what is being taught in this program.	Before	967	3.91	p<0.001**	3%	5%	23%	37%	32%
	Now	960	4.35		2%	2%	9%	32%	55%
3. I like what I am learning in this program.	Before	956	3.66	p<0.001**	3%	7%	34%	32%	24%
	Now	948	4.09		4%	3%	16%	34%	43%
4. I think I will be able to use what I learn in this program in other classes.	Before	959	3.72	p<0.001**	4%	7%	28%	37%	25%
	Now	950	4.19		3%	3%	13%	36%	45%
5. Even when I do poorly on a test, I try to learn from my mistakes.	Before	965	3.93	p<0.001**	3%	6%	19%	39%	33%
	Now	959	4.35		2%	2%	9%	36%	52%
6. I think that what I am learning in this program is useful for me to know.	Before	960	3.75	p<0.001**	4%	7%	28%	32%	29%
	Now	955	4.20		3%	3%	13%	32%	49%
7. I think that what we are learning in this program is interesting.	Before	964	3.58	p<0.001**	6%	9%	31%	31%	24%
	Now	959	4.01		5%	4%	18%	32%	41%
8. Understanding STEM (Science, Technology, Engineering, and Math) is important to me.	Before	963	3.67	p<0.001**	5%	8%	29%	29%	28%
	Now	958	4.18		3%	3%	14%	33%	47%
9. I enjoy STEM (Science, Technology, Engineering, and Math) in general.	Before	965	3.56	p<0.001**	6%	11%	31%	26%	27%
	Now	959	4.00		5%	6%	18%	30%	42%

Note. [1] Reference lines are set at 3.5 and 4. [2] Please note that only students with matched Pre and Post data were assessed for significance. Desired statistically significant changes are highlighted in green and undesired statistically significant changes are highlighted in red. **p<0.001, *p<0.01, †p<0.05. Highest percentages are highlighted in gray.

Table 5. Self-Regulation/Self-Motivation

Self-Regulation/Self-Motivation		n	Mean[1]	Paired Samples t-test[2]	1 (Strongly Disagree)	2 (Disagree)	3 (Neutral)	4 (Agree)	5 (Strongly Agree)
10. I turn all my assignments in on time.	Before	966	3.69	p<0.001**	3%	10%	30%	32%	26%
	Now	959	3.93		1%	6%	24%	37%	32%
11. I miss class often. (n)	Before	961	1.63	p=0.147	65%	18%	9%	5%	3%
	Now	954	1.59		69%	15%	7%	5%	3%
12. I am often late for class. (n)	Before	952	1.61	p=0.828	64%	20%	10%	5%	2%
	Now	948	1.60		67%	18%	7%	4%	4%
13. I set aside time to do my homework and study.	Before	960	3.33	p<0.001**	8%	14%	32%	29%	17%
	Now	959	3.66		6%	7%	27%	34%	26%
14. When I say I'm going to do something, I do it.	Before	964	3.82	p<0.001**	2%	7%	26%	38%	28%
	Now	962	4.07		1%	4%	19%	38%	38%
15. I am a hard worker.	Before	961	4.11	p<0.001**	2%	4%	18%	33%	43%
	Now	961	4.34		1%	1%	13%	34%	52%
16. I finish whatever I begin.	Before	958	3.89	p<0.001**	1%	6%	27%	34%	32%
	Now	956	4.16		1%	2%	20%	34%	43%

Note. [1] Reference lines are set at 3.5 and 4. [2] Please note that only students with matched Pre and Post data were assessed for significance. Desired statistically significant changes are highlighted in green and undesired statistically significant changes are highlighted in red. **p<0.001, *p<0.01, †p<0.05; (n) negatively worded statement. Highest percentages are highlighted in gray.

Table 6. Intent to Persist

Intent to Persist		n	Mean[1]	Paired Samples t-test[2]	1 (Strongly Disagree)	2 (Disagree)	3 (Neutral)	4 (Agree)	5 (Strongly Agree)
17. I am considering a career in STEM (Science, Technology, Engineering, and Math).	Before	963	3.28	p<0.001**	12%	14%	31%	21%	23%
	Now	962	3.65		9%	9%	24%	25%	33%
18. I intend to get a college degree in STEM (Science, Technology, Engineering, and Math).	Before	964	3.36	p<0.001**	10%	15%	30%	22%	24%
	Now	960	3.70		8%	10%	22%	26%	35%
19. I can see myself working in STEM (Science, Technology, Engineering, and Math).	Before	960	3.31	p<0.001**	11%	14%	31%	22%	22%
	Now	959	3.65		9%	9%	23%	27%	32%
20. Someday, I would like to have a career in STEM (Science, Technology, Engineering, and Math).	Before	957	3.29	p<0.001**	12%	14%	32%	20%	23%
	Now	957	3.59		10%	10%	24%	25%	32%
21. I intend to graduate from high school.	Before	958	4.72	p<0.001**	2%	1%	5%	8%	84%
	Now	959	4.83		1%	0%	4%	6%	89%

Note. [1] Reference lines are set at 3.5 and 4. [2] Please note that only students with matched Pre and Post data were assessed for significance. Desired statistically significant changes are highlighted in green and undesired statistically significant changes are highlighted in red. **p<0.001, *p<0.01, †p<0.05. Highest percentages are highlighted in gray.

Table 7. Problem Solving, Now Only

Problem Solving	n	Mean[1]	Assessment	1 (Strongly Disagree)	2 (Disagree)	3 (Neutral)	4 (Agree)	5 (Strongly Agree)
22. In this program, my teacher(s) tells me how to improve my work.	927	4.10	Good ☺	3%	4%	16%	33%	44%
23. In this program, my teacher(s) lets us choose our own topics or projects to investigate.	912	3.50	Attention ✓	8%	11%	28%	26%	26%
24. In this program, I work out explanations on my own.	957	3.89	Attention ✓	1%	3%	26%	45%	24%
25. In this program, I have opportunities to explain my ideas.	958	3.90	Attention ✓	3%	5%	20%	42%	30%
26. In this program, we plan and do our own projects and/or experiments.	959	3.83	Attention ✓	4%	6%	23%	38%	29%
27. In this program, we work on real-world problems.	961	4.02	Good ☺	3%	5%	18%	36%	39%
28. In this program, we have class discussions.	957	4.20	Good ☺	2%	2%	13%	39%	44%
29. In this program, we investigate to see if our ideas are right.	955	4.03	Good ☺	2%	4%	18%	40%	36%
30. In this program, we need to be able to think and ask questions.	953	4.28	Good ☺	2%	2%	11%	38%	47%
31. In this program, we are expected to understand and explain ideas.	953	4.27	Good ☺	2%	1%	11%	39%	47%

Note. [1] Reference lines are set at 3.5 and 4. Assessment: Good=Above 4.0; Attention=Below 4.0; Action=Below 3.5. Highest percentages are highlighted in gray.

Table 8. Implementation Activities, Now Only

Implementation Activities	n	Mean[1]	Assessment	1 (Strongly Disagree)	2 (Disagree)	3 (Neutral)	4 (Agree)	5 (Strongly Agree)
32. In this program, my teacher(s) takes notice of students' ideas.	936	3.87	Attention ✓	4%	6%	22%	36%	33%
33. In this program, my teacher(s) shows us how new information relates to what we have already learned.	924	4.10	Good ☺	3%	2%	16%	38%	41%
34. In this program, we learn what scientists/ technicians/ engineers/ mathematicians or other STEM professionals do.	956	3.88	Attention ✓	5%	6%	17%	37%	34%
35. In this program, we do our work in groups.	953	3.99	Attention ✓	2%	2%	27%	36%	34%
36. In this program, we interact with scientists/ technicians/ engineers/ mathematicians or other STEM professionals.	957	3.68	Attention ✓	8%	8%	22%	32%	30%

Note. [1] Reference lines are set at 3.5 and 4. Assessment: Good=Above 4.0; Attention=Below 4.0; Action=Below 3.5. Highest percentages are highlighted in gray.

Table 9. Educational Plans

What is the highest level of education you plan to achieve?	Before		Now		*Change*[1]	
	n	%	n	%	*n*	*%*
High School	164	17%	89	10%	-75	-7.88%
2-year college	121	13%	71	8%	-50	-5.24%
4-year college	288	31%	214	23%	-74	-7.63%
Graduate School	185	20%	236	25%	+51	+5.70%
Professional School	182	19%	320	34%	+138	+15.05%
Total	**940**	**100%**	**930**	**100%**		
Average[2]	**2.91**		**3.33**		**p<0.001**(significant)[3]**	

Note. [1] Change from Before to Now. Increases are highlighted in green; decreases are highlighted in red.
[2]To compute averages, the following codes were applied: High School (1), 2-year college (2), 4-year college (3), Graduate School (4), Professional School (4).
[3]Paired samples t-test, p-value: **p<0.001, *p<0.01, †p<0.05.

Table 10. Demographics

Gender	n	%
Female	422	44%
Male	531	56%
Total	**954**	**100%**

Ethnicity	n	%
Asian	27	3%
Black	362	38%
Hispanic	89	9%
Native American	3	0%
White	371	39%
Multiracial	65	7%
Other	36	4%
Total	**953**	**100%**

Grade	n	%
6th	115	12%
7th	179	19%
8th	303	32%
9th	14	1%
10th	72	7%
11th	152	16%
12th	125	13%
Other	1	0%
Total	**961**	**100%**

Table 11. Participation

How long have you participated in this program?		n	%
Dosage	0 semesters	36	4%
	1 semester	161	17%
	2 semesters	402	42%
	3 semesters	64	7%
	4 or more semesters	239	25%
	Summer Only	1	0%
	Don’t Know	53	6%
	Total	**956**	**100%**
Did you participate in this program during the summer?		**n**	**%**
Summer Participation	No	645	68%
	Yes	233	25%
	Don't Know	71	7%
	Total	**949**	**100%**

Table 12. Program Rating

Program Rating: How would you rate this program?	**n**	**Mean[1]**	**Assessment**	**1 (Very Poor)**	**2 (Poor)**	**3 (Average)**	**4 (Good)**	**5 (Excellent)**
	949	4.20	Good ☺	3%	2%	13%	36%	45%

Note. [1] Reference lines are set at 3.5 and 4. Assessment: Good=Above 4.0; Attention=Below 4.0; Action=Below 3.5. Highest percentage is highlighted in gray.

Appendix A. Construct Reliabilities

Table A1. Construct Reliabilities (n=970)

Constructs		Cronbach's alpha	*Reliability Interpretation*
Intrinsic Motivation (9-items)	Before	0.888	*Very good*
	Now	0.912	*Excellent*
Self-Management/Self-Regulation (7-items)	Before	0.710	*Good*
	Now	0.739	*Good*
Intent to Persist (5-items)	Before	0.887	*Very good*
	Now	0.891	*Very good*
Problem Solving (10-items)	Now	0.897	*Very good*
Implementation Activities (5-items)	Now	0.848	*Very good*

Cronbach's Alpha Reliability Key: Cronbach's alpha is a measure of the internal consistency of items in a construct. This statistic ranges from 0 to 1.00; the higher the value the better. An alpha of .80 or higher is considered to have achieved very good measurement reliability; an alpha of .65 is considered acceptable (Field, 2009).

Reliability	Interpretation
.90 and above	Excellent reliability; at the level of the best measures
.80 - .90	Very good
.70 - .80	Good; in the range of most. There are probably a few items which could be improved.
.60 - .70	Somewhat low. This measure needs to be supplemented by other measures (e.g., more surveys) to determine outcomes. There are probably some items which could be improved.
.50 - .60	Suggests need for revision of measure, unless it is quite short (ten or fewer items). The test definitely needs to be supplemented by other measures (e.g., more tests).
.50 or below	Questionable reliability. This measure should not contribute heavily to the outcomes and needs revision.

From: J. C. Nunnally, Psychometric Theory. New York: McGraw-Hill, 1967, pp. 172-235.

Reference:

Field, A. (2009). *Discovering Statistics Using SPSS, 3rd Edition*. Thousand Oaks, CA: Sage Publications.

Appendix B. Disaggregated Findings by Dosage

Evaluators from the Real STEM Georgia Southern program informed SageFox Consulting Group in May 2015 that their program consisted of varying treatment conditions. Given the differences in program dosage, SageFox provided this program with disaggregated findings for each treatment group. A summary of the disaggregated data for each treatment/dosage condition is described in detail below. This information should be used in conjunction with the overall programmatic data displayed in the Executive Summary to inform any modifications or suggestions for improvement.

Real STEM Georgia Southern University

The Real STEM partnership program with Georgia Southern University consisted of three treatment levels:

- *Treatment 1- Full Scientific Research Course*: high schools offering a full research course;
- *Treatment 2- Module/unit only- Second Year*: middle schools offering a unit for the second time; and,
- *Treatment 3-Module/unit only-First Year*: middle and high schools offering a unit for the first time.

Given the differences in duration for each treatment group described above, the current analysis displays separate findings for each treatment level. In particular, the following numbers of students were included per treatment level:

School-Teacher	Treatment 1	Treatment 2	Treatment 3
Statesboro High School- Rich McCombs	14	--	--
Burke County High School- Justin Russell	13	--	--
Camden County High School- Theresa Lyster	20	--	--
Brantley County Middle School- Beth Ann Thomas	--	15	--
Brantley County Middle School- Danielle Lopez	--	5	--
Brantley County Middle School- Debra Deems	--	11	--
Brantley County Middle School- Deon Horne	--	16	--
Brantley County Middle School- Gary Edholm	--	17	--
Brantley County Middle School- Grace MacMillan	--	16	--
Brantley County Middle School- John Smith	--	7	--
Brantley County Middle School- Lois Hendrix	--	18	--
Richmond Hill Middle School- John Melcher	--	--	63
Total	**47**	**105**	**63**

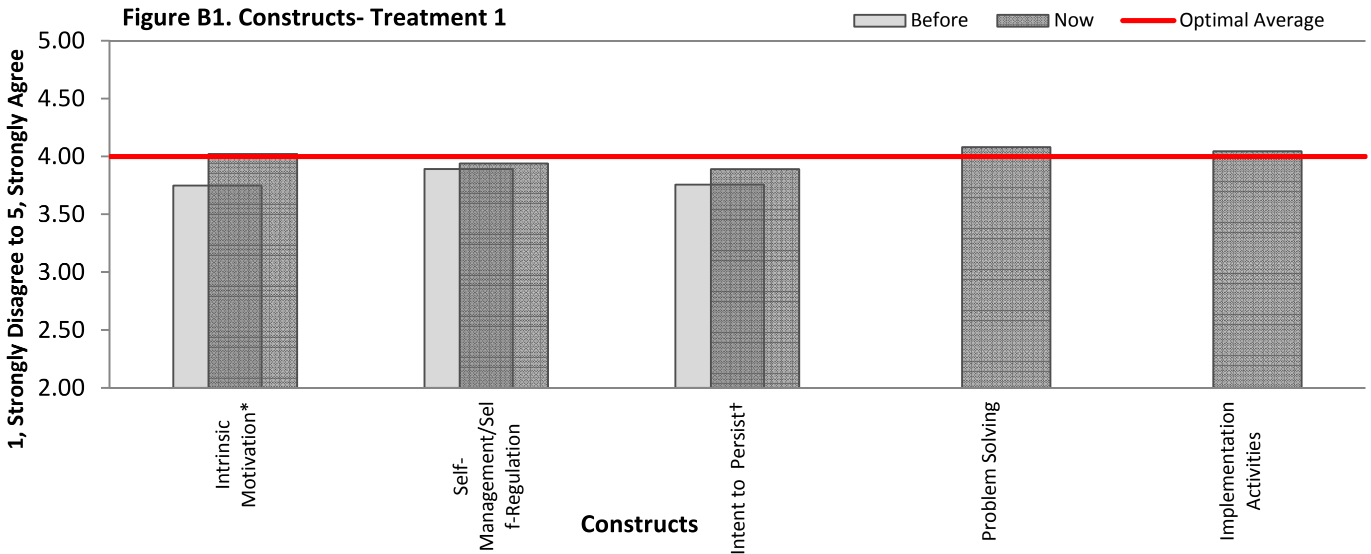
Table B1 summarizes students' responses per treatment level. Among students in Treatment 1 (e.g., high schools offering a full research course), statistically significant increases were detected for two constructs: *Intrinsic Motivation* and *Intent to Persist*. That is, from before the program to now, students in the full research course show a statistically significant increase in their motivation to learn the material and to pursue an education and career in STEM. Students in Treatment 2 (e.g., middle school schools offering a unit for the second time) and Treatment 3 (e.g., middle and high schools offering a unit for the first time) reported statistically significant increases across all constructs from before the program to now: *Intrinsic Motivation, Self-Management/Self-Regulation,* and *Intent to Persist.* Examining students' 'now' scores, it is evident that Treatment 3 achieved the highest means across all constructs; by contrast, students in Treatment 2 show the lowest 'now' score averages across most constructs. In fact, most construct averages for Treatment 2 students did not reach or exceed the optimal average of 4.0 on a 5-point Likert scale (1, Strongly Disagree to 5, Strongly Agree). This may suggest that additional attention is needed to enhance the inquiry-based learning environment in those classrooms offering the unit for a second time (Treatment 2).

Table B1. Summary of Results by Constructs

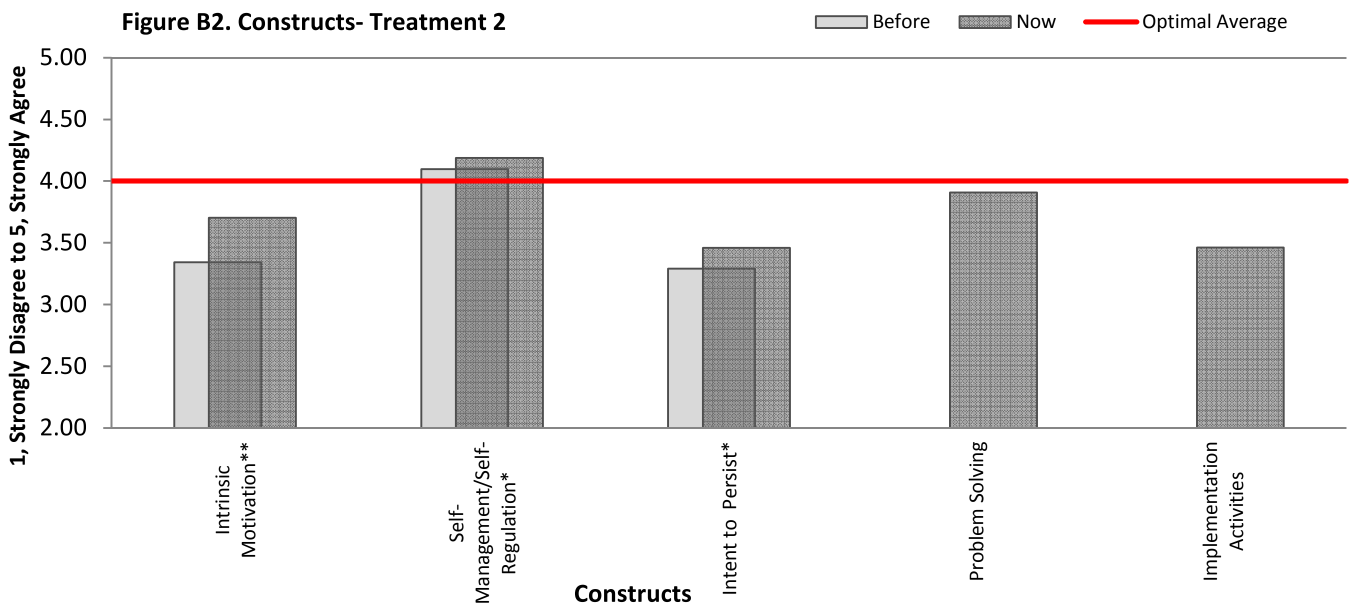
Overall- Constructs										
		Treatment 1: Full Scientific Research Course			Treatment 2: Module/unit only – Second year			Treatment 3: Module/unit only- First year		
Constructs		**n**	**Mean**	**Paired Samples t-test**	**n**	**Mean**	**Paired Samples t-test**	**n**	**Mean**	**Paired Samples t-test**
Intrinsic Motivation	Before	47	3.75	p=0.007*	105	3.34	p<0.001**	63	3.93	p<0.001**
	Now	47	4.02		104	3.70		62	4.71	
Self-Management/ Self-Regulation	Before	47	3.89	p=0.168	104	4.10	p=0.004*	63	4.22	p<0.001**
	Now	47	3.94		105	4.19		63	4.55	
Intent to Persist	Before	47	3.76	p=0.029†	103	3.29	p=0.001*	63	3.87	p<0.001**
	Now	47	3.89		105	3.46		63	4.39	
Problem Solving	Now	47	4.08	N/A	105	3.91	N/A	63	4.64	N/A
Implementation Activities	Now	47	4.04	N/A	104	3.46	N/A	62	4.60	N/A

Note. Please note that only students with matched Pre and Post data were assessed for significance. Desired statistically significant changes are highlighted in green and undesired statistically significant changes are highlighted in red. **p<0.001, *p<0.01, †p<0.05. Negatively worded statements were reverse coded for mean computations.

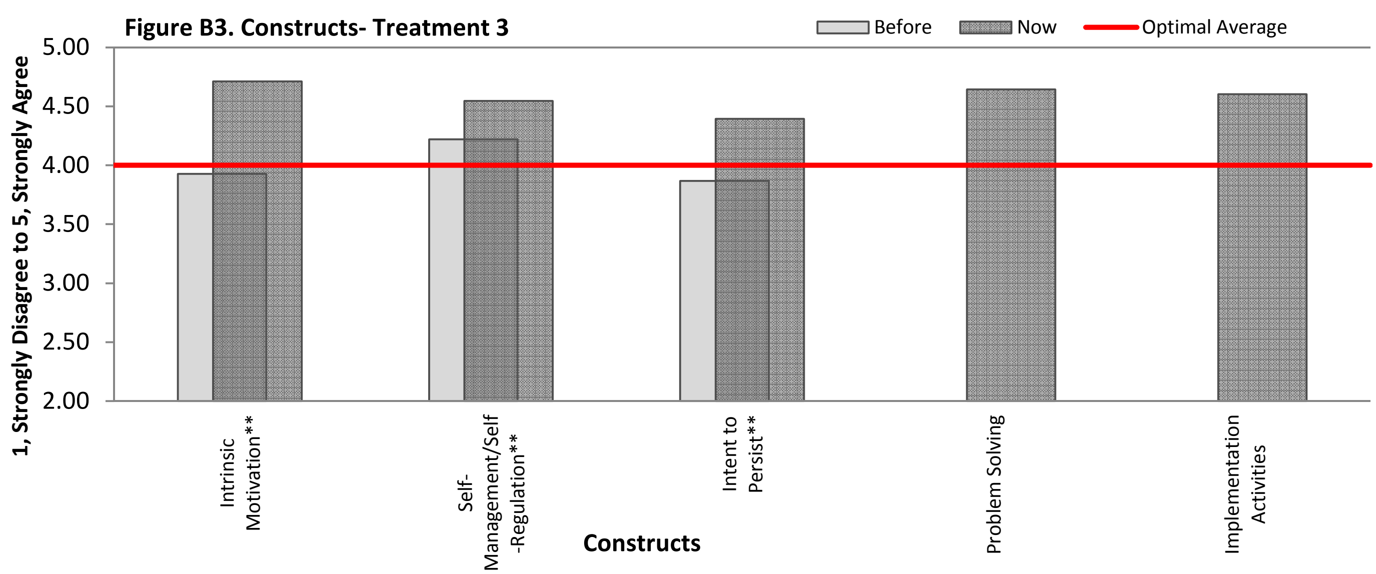
Figure B1. Constructs- Treatment 1



Note. A paired samples t-test was used to find the p-value.**p<0.001, *p<0.01, †p<0.05; Scale is truncated for visual clarity.



Note. A paired samples t-test was used to find the p-value.**p<0.001, *p<0.01, †p<0.05; Scale is truncated for visual clarity.



Note. A paired samples t-test was used to find the p-value. **p<0.001, *p<0.01, †p<0.05; Scale is truncated for visual clarity.