APPLIED LEARNING STUDENT QUESTIONNAIRE: *OVERALL ANALYSIS*

Overall Results
December 2014

Executive Summary

Participants and Methods

In December 2014, 1,042 students across 6 Race to the Top programs completed the Applied Learning Student Questionnaire (ALSQ). The response rates displayed in Table 1 suggest that 84% of the total number of participating students responded to the survey. The response rates per program ranged from 33% to 100%. 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	314	341	92%
Real STEM Georgia Southern	212	266	80%
RT3 Computational Thinking	11	33	33%
STEM for Life Carroll County	199	300	66%
STEP Academy Gwinnett	230	230	100%
Tift County Mechatronics	76	76	100%
Total	1,042	1,246	84%

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.

Prepared by:
sengelman@sagefoxgroup.com

February 2014
tmcklin@sagefoxgroup.com

SageFox CONSULTING GROUP
30 Boltwood Walk • Amherst, Mass. 01002
413.256.6169 • www.sagefoxgroup.com

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* and *Intent to Persist*; a small effect size was found for *Self-Management/Self-Regulation*. Across all constructs, the largest effect size observed was for the *Intrinsic Motivation* construct (*d*=0.76). 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 55% of students said that understanding STEM is important to them compared to 81% 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 the following 3 constructs— *Intent to Persist*, *Problem Solving*, 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.

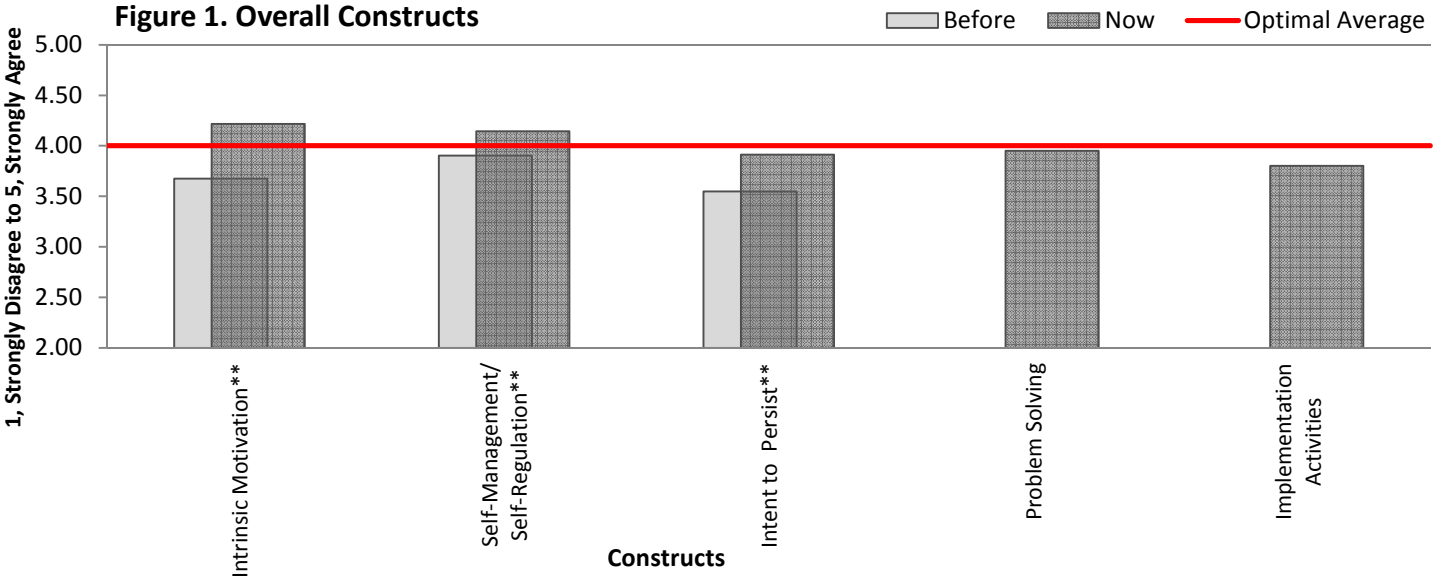
Table 2. Summary of Results by Constructs

Overall- Constructs						
Constructs		**n**	**Mean[1]**		**Paired Samples t-test[2]**	**Effect Size (interpretation)[3]**
Intrinsic Motivation	Before	1040		3.68	p<0.001**	*d*=0.76 (Medium)
	Now	1028		4.22		
Self-Management/Self-Regulation	Before	1035		3.90	p<0.001**	*d*=0.45 (Small)
	Now	1030		4.14		
Intent to Persist	Before	1032		3.55	p<0.001**	*d*=0.54 (Medium)
	Now	1030		3.91		
Problem Solving	Now	1031		3.95	N/A	N/A
Implementation Activities	Now	1026		3.80	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 (<.5); Medium (.5 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.

Figure 1. Overall Constructs



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

- **ALSQ Survey Constructs by Program**

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*. It is important to note that due to the small sample size (n=11) for the RT3 Computational Thinking program, statistical power[4] was compromised. That is, a smaller sample size decreases the chance of finding a significant difference. Thus, the lack of statistically significant findings for RT3 Computational Thinking may be due to the small sample size and not the program intervention. Examining effect sizes, students in the following three programs show medium to large effect sizes: Tift County Mechatronics, STEP Academy Gwinnett, and Real STEM Georgia Southern University. This suggests that the above mentioned programs had a medium to large impact on students' attitudes.[5]

Table 3. Summary of Results by Constructs per Program

Overall- Constructs per Program										
Constructs		**21st Century Rockdale County (n=314)**			**Real STEM Georgia Southern (n=212)**			**RT3 Computational Thinking (n=11)**		
		Mean	*t-test*	Effect Size	Mean	*t-test*	Effect Size	Mean	*t-test*	Effect Size
Intrinsic Motivation	Before	3.85	p<0.001**	0.59 (M)	3.60	p<0.001**	1.04 (L)	3.65	p=0.030†	0.76 (M)
	Now	4.16			4.33			3.25		
Self-Management/ Self-Regulation	Before	4.07	p<0.001**	0.26 (S)	4.05	p<0.001**	0.54 (M)	4.55	p=0.136	0.49 (S)
	Now	4.17			4.27			4.47		
Intent to Persist	Before	3.70	p<0.001**	0.43 (S)	3.58	p<0.001**	0.59 (M)	4.22	p=1.000	0.00
	Now	3.91			3.96			4.22		
Problem Solving	Now	3.82	N/A	N/A	4.31	N/A	N/A	3.25	N/A	N/A
Implementation Activities	Now	3.58			4.14			3.13		

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 (S) (<.5); Medium (M) (.5 to .8); Large (L) (>.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] 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 the effect size (the salience of the treatment relative to the noise in measurement).

[5] For additional information related to 2 programs (STEM for Life Carroll County and Real STEM Georgia Southern University) see Appendix B.

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=199)**			**STEP Academy Gwinnett (n=230)**			**Tift County Mechatronics (n=76)**		
		Mean	*t-test*	Effect Size	Mean	*t-test*	Effect Size	Mean	*t-test*	Effect Size
Intrinsic Motivation	Before	3.54	p<0.001**	0.63 (M)	3.53	p<0.001**	0.87 (L)	3.97	p<0.001**	1.22 (L)
	Now	4.06			4.19			4.72		
Self-Management/ Self-Regulation	Before	3.74	p<0.001**	0.46 (S)	3.57	p<0.001**	0.58 (M)	4.13	p<0.001**	0.65 (M)
	Now	4.00			3.99			4.46		
Intent to Persist	Before	3.36	p<0.001**	0.45 (S)	3.37	p<0.001**	0.64 (M)	3.81	p<0.001**	1.01 (L)
	Now	3.71			3.80			4.60		
Problem Solving	Now	3.78	N/A	N/A	3.78	N/A	N/A	4.56	N/A	N/A
Implementation Activities	Now	3.72			3.60			4.65		

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 (S) (<.5); Medium (M) (.5 to .8); Large (L) (>.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 5 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 *Problem Solving* and *Implementation Activities*.

Figure 2. Intrinsic Motivation ("Now" Scores)

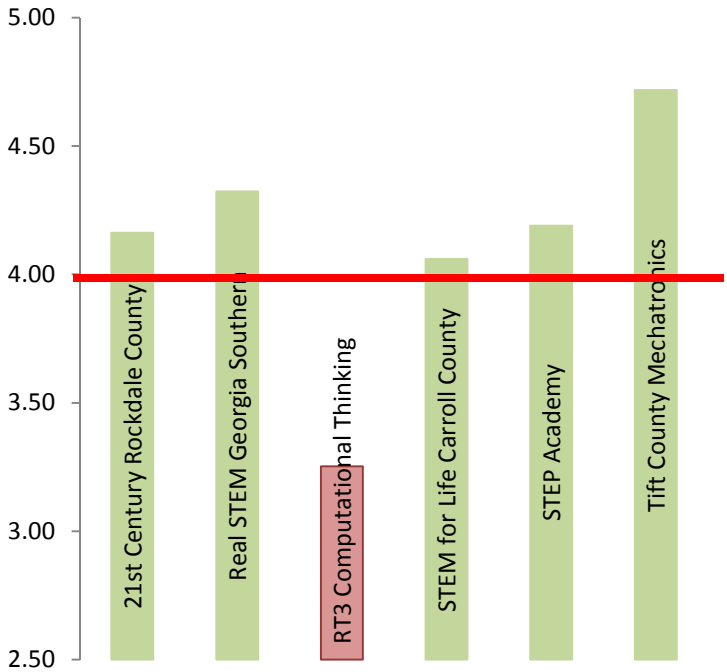
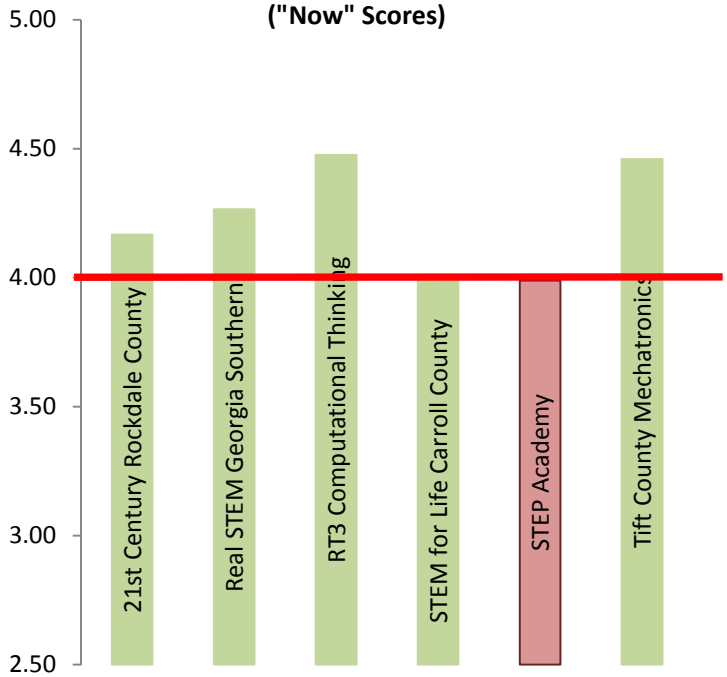


Figure 3. Self-Management/Self-Regulation ("Now" Scores)



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.

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

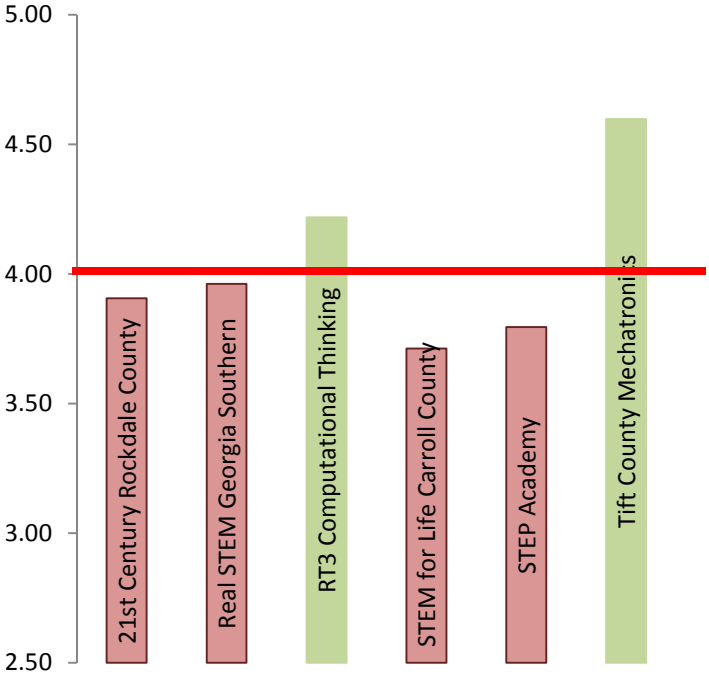


Figure 5. Problem Solving ("Now" Scores)

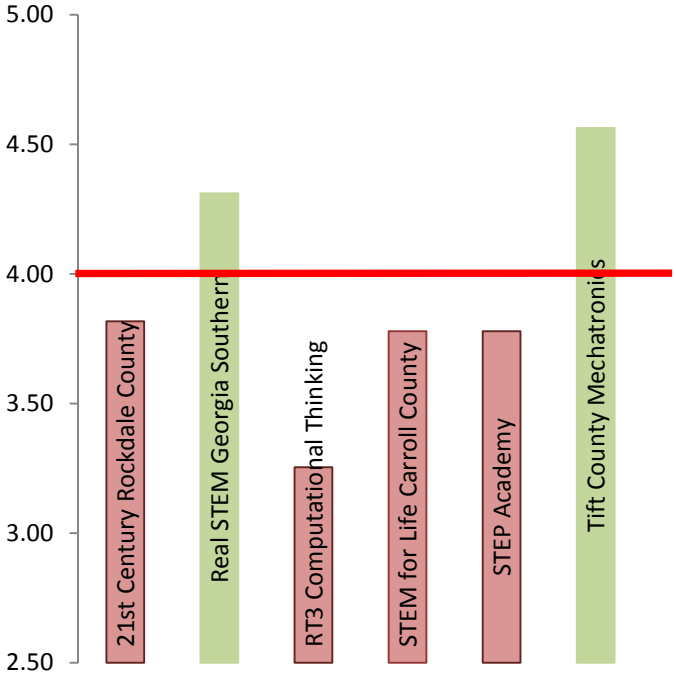


Figure 6. Implementation Activities ("Now" Scores)

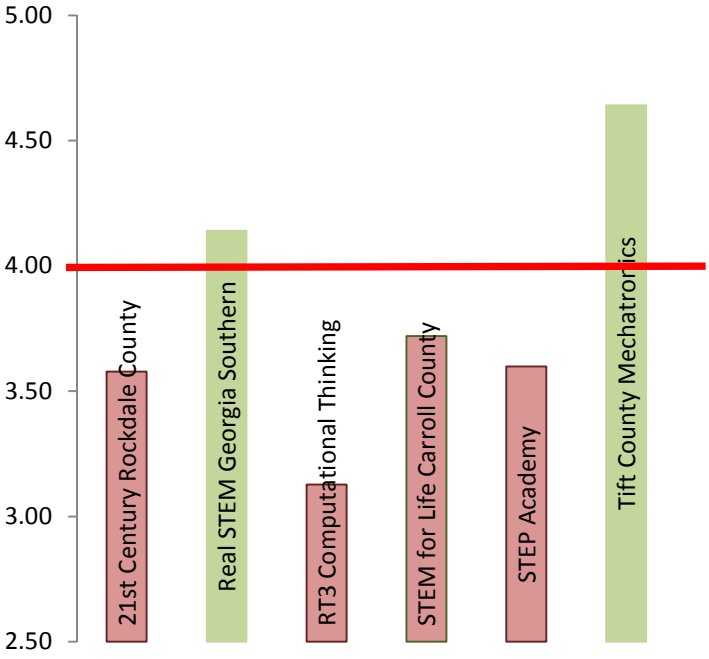
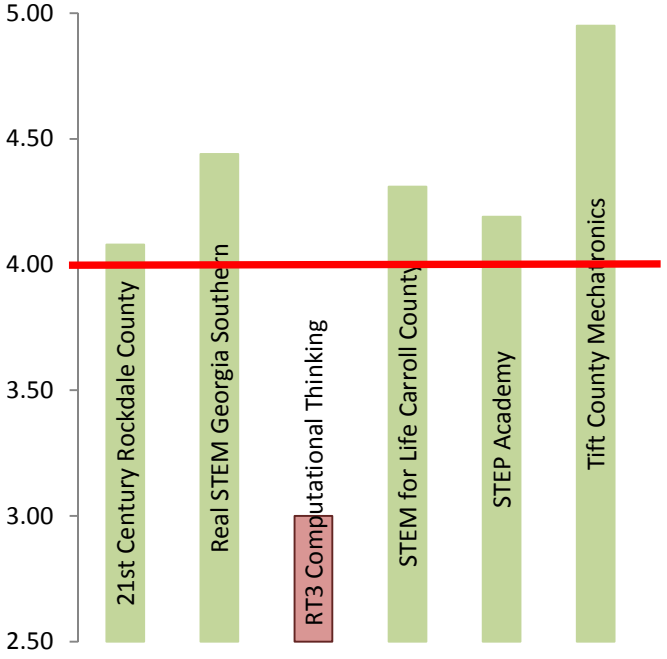


Figure 7. Overall Program Ratings



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 4.27. See Table 12. Looking at Figure 7, it is evident that 5 out of 6 programs were rated above the optimal average. The RT3 Computational Thinking program may need additional assistance in improving student enjoyment.

- **Areas for Further Improvement**
  Across all programs, further enhancing problem solving skills, 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, and plan and conduct their own projects. Likewise, encouraging programs to provide activities that foster interaction with STEM professionals may increase student exposure to real-world applications and careers. Such implementation activities may strengthen students' intentions and motivations to pursue additional education in STEM fields.

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	1040	3.33		p<0.001**		7%	13%	35%	30%	15%
	Now	1026	3.92				3%	5%	21%	39%	32%
2. It is important to me to learn what is being taught in this program.	Before	1036	3.97		p<0.001**		2%	5%	21%	40%	33%
	Now	1028	4.43				1%	1%	8%	31%	58%
3. I like what I am learning in this program.	Before	1029	3.68		p<0.001**		3%	7%	31%	37%	22%
	Now	1027	4.20				2%	4%	13%	34%	47%
4. I think I will be able to use what I learn in this program in other classes.	Before	1030	3.65		p<0.001**		4%	9%	30%	35%	23%
	Now	1026	4.24				1%	3%	12%	35%	48%
5. Even when I do poorly on a test, I try to learn from my mistakes.	Before	1033	3.93		p<0.001**		3%	6%	19%	39%	33%
	Now	1028	4.45				0%	2%	7%	34%	57%
6. I think that what I am learning in this program is useful for me to know.	Before	1023	3.78		p<0.001**		2%	7%	28%	37%	26%
	Now	1018	4.32				1%	2%	12%	31%	53%
7. I think that what we are learning in this program is interesting.	Before	1022	3.59		p<0.001**		3%	9%	35%	32%	21%
	Now	1017	4.08				2%	4%	19%	34%	41%
8. Understanding STEM (Science, Technology, Engineering, and Math) is important to me.	Before	1029	3.63		p<0.001**		4%	9%	31%	30%	25%
	Now	1022	4.23				2%	3%	15%	33%	48%
9. I enjoy STEM (Science, Technology, Engineering, and Math) in general.	Before	1020	3.51		p<0.001**		5%	10%	37%	28%	21%
	Now	1025	4.06				3%	5%	17%	35%	41%

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	1035	3.65	p<0.001**	3%	11%	29%	32%	25%
	Now	1025	3.96		2%	5%	21%	37%	34%
11. I miss class often. (n)	Before	1033	1.70	p=0.171	61%	21%	10%	5%	4%
	Now	1023	1.66		65%	17%	8%	4%	5%
12. I am often late for class. (n)	Before	1008	1.66	p=0.939	61%	21%	10%	5%	2%
	Now	1006	1.66		65%	18%	8%	5%	4%
13. I set aside time to do my homework and study.	Before	1031	3.38	p<0.001**	6%	14%	32%	30%	17%
	Now	1030	3.80		4%	6%	26%	35%	30%
14. When I say I'm going to do something, I do it.	Before	1032	3.74	p<0.001**	2%	8%	29%	35%	26%
	Now	1030	4.07		2%	2%	21%	39%	37%
15. I am a hard worker.	Before	1031	4.08	p<0.001**	1%	5%	18%	36%	39%
	Now	1028	4.36		1%	1%	11%	34%	53%
16. I finish whatever I begin.	Before	1026	3.83	p<0.001**	2%	6%	27%	35%	30%
	Now	1029	4.13		2%	1%	20%	37%	40%

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	1032		3.25	p<0.001**		12%	14%	31%	21%	22%
	Now	1030		3.69			9%	8%	23%	24%	35%
18. I intend to get a college degree in STEM (Science, Technology, Engineering, and Math).	Before	1030		3.36	p<0.001**		9%	13%	33%	22%	22%
	Now	1027		3.77			7%	7%	24%	24%	37%
19. I can see myself working in STEM (Science, Technology, Engineering, and Math).	Before	1029		3.22	p<0.001**		11%	15%	33%	22%	19%
	Now	1029		3.64			8%	9%	25%	26%	31%
20. Someday, I would like to have a career in STEM (Science, Technology, Engineering, and Math).	Before	1029		3.22	p<0.001**		11%	15%	33%	21%	19%
	Now	1020		3.64			8%	9%	25%	26%	32%
21. I intend to graduate from high school.	Before	1028		4.69	p<0.001**		1%	2%	5%	10%	82%
	Now	1030		4.81			1%	0%	3%	7%	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.	1025		4.06	Good ☺		2%	5%	15%	37%	40%
23. In this program, my teacher(s) lets us choose our own topics or projects to investigate.	1014		3.43	Action !		7%	13%	32%	27%	21%
24. In this program, I work out explanations on my own.	1031		3.76	Attention ✓		1%	4%	32%	45%	18%
25. In this program, I have opportunities to explain my ideas.	1031		3.86	Attention ✓		2%	6%	23%	42%	27%
26. In this program, we plan and do our own projects and/or experiments.	1028		3.76	Attention ✓		4%	7%	27%	37%	26%
27. In this program, we work on real-world problems.	1030		3.96	Attention ✓		3%	5%	20%	39%	33%
28. In this program, we have class discussions.	1028		4.17	Good ☺		2%	3%	14%	38%	43%
29. In this program, we investigate to see if our ideas are right.	1026		4.00	Good ☺		2%	4%	20%	41%	33%
30. In this program, we need to be able to think and ask questions.	1026		4.28	Good ☺		1%	1%	13%	38%	47%
31. In this program, we are expected to understand and explain ideas.	1027		4.23	Good ☺		1%	1%	14%	39%	44%

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.	1026		3.82	Attention ✓		4%	5%	24%	36%	30%
33. In this program, my teacher(s) shows us how new information relates to what we have already learned.	1015		4.10	Good ☺		2%	4%	15%	39%	39%
34. In this program, we learn what scientists/ technicians/ engineers/ mathematicians or other STEM professionals do.	1025		3.74	Attention ✓		5%	6%	25%	38%	26%
35. In this program, we do our work in groups.	1020		3.82	Attention ✓		3%	4%	30%	33%	30%
36. In this program, we interact with scientists/ technicians/ engineers/ mathematicians or other STEM professionals.	1023		3.53	Attention ✓		7%	11%	26%	35%	21%

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	194	19%	108	11%	-86	-8.41%
2-year college	130	13%	77	8%	-53	-5.16%
4-year college	289	29%	216	22%	-73	-6.94%
Graduate School	215	21%	276	28%	+61	+6.53%
Professional School	174	17%	309	31%	+135	+13.97%
Total	**1002**	**100%**	**986**	**100%**		
Average[2]	**2.87**		**3.30**		**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	489	48%
Male	534	52%
Total	**1023**	**100%**

Ethnicity	n	%
Asian	42	4%
Black	424	41%
Hispanic	147	14%
Native American	3	0%
White	313	31%
Multiracial	63	6%
Other	32	3%
Total	**1024**	**100%**

Grade	n	%
6th	152	15%
7th	145	14%
8th	354	35%
9th	5	0%
10th	49	5%
11th	124	12%
12th	188	18%
Other	8	1%
Total	**1025**	**100%**

Table 11. Participation

How long have you participated in this program?		n	%
Dosage	0 semesters	6	0%
	1 semester	143	11%
	2 semesters	615	48%
	3 semesters	41	3%
	4 or more semesters	296	23%
	Summer Only	2	0%
	Don't Know	167	13%
	Total	**1270**	**100%**

Did you participate in this program during the summer?		n	%
Summer Participation	No	891	70%
	Yes	235	19%
	Don't Know	143	11%
	Total	**1270**	**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)
	1023	4.27	Good ☺		1%	2%	14%	34%	49%

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

Construct Reliabilities				
Constructs		**n**	**Cronbach's alpha**	***Reliability Interpretation***
Intrinsic Motivation (9-items)	Before	978	0.881	*Very good*
	Now	964	0.900	*Excellent*
Self-Management/Self-Regulation (7-items)	Before	985	0.603	*Somewhat Low*
	Now	968	0.656	*Somewhat Low*
Intent to Persist (5-items)	Before	1017	0.871	*Very good*
	Now	1006	0.882	*Very good*
Problem Solving (10-items)	Now	977	0.893	*Very good*
Implementation Activities (5-items)	Now	996	0.844	*Very good*

Note. Construct reliabilities were computed based on December 2014 data. The sample size displayed (n) reflects the number of students who answered all items related to each construct.

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 two programs—Real STEM Georgia Southern University and STEM for Life Carroll County— informed SageFox Consulting Group in December 2014 that their programs consisted of varying treatment or dosage conditions. Given the differences in program dosage, SageFox provided these programs with disaggregated findings for each condition. 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.

1. 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[6]:

School-Teacher	Treatment 1	Treatment 2	Treatment 3
Statesboro High School- Rich McCombs	12	--	--
Burke County High School- Justin Russell	6	--	--
William James Middle School- Amy Smith	--	--	18
Richmond Hill Middle School- John Melcher	--	--	95
Statesboro High School- Lee Bratton	--	--	64
Langston Chapel Middle School- Broucek	--	--	17
Total	**18**	**--**	**194**

Table B1 summarizes students' responses per treatment level. Among students in Treatment 3 (e.g., middle and high schools offering a unit for the first time), statistically significant increases were detected across all constructs from before the program to now: *Intrinsic Motivation*, *Self-Management/Self-Regulation Skills* and *Intent to Persist*. The largest student gains observed were in the *Intrinsic Motivation* construct. Before the program, students in Treatment 3 rated their motivation to learn about STEM at a mean of 3.60 on a 5-point Likert scale (1, *Strongly Disagree* to 5, *Strongly Agree*); now, students indicate that they are intrinsically motivated to engage in STEM-related tasks and projects with a mean of 4.34 on a 5-point Likert scale. Among students in Treatment 1 (e.g., high schools offering a full research course), statistically significant increases were detected for only one construct: *Intrinsic Motivation*. However, it is important to note that due to the small sample size (n=18) statistical power[7] was compromised. That is, a smaller sample size decreases the chance of finding a significant difference. Thus, the lack of statistically significant findings for Treatment 1 may be due to the small sample size and not the program intervention. Examining the "now" scores, it is evident that all constructs, with the exception of *Intent to Persist*, reached or exceeded the optimal average of 4.00 (1, *Strongly Disagree* to 5, *Strongly Agree*) for students across Treatment 1 and Treatment 3. See

[6] Students participating in Treatment 2 will be completing the ALSQ in Spring 2015 only; thus, only students participating in Treatments 1 and 3 who completed the ALSQ are included in the current report.

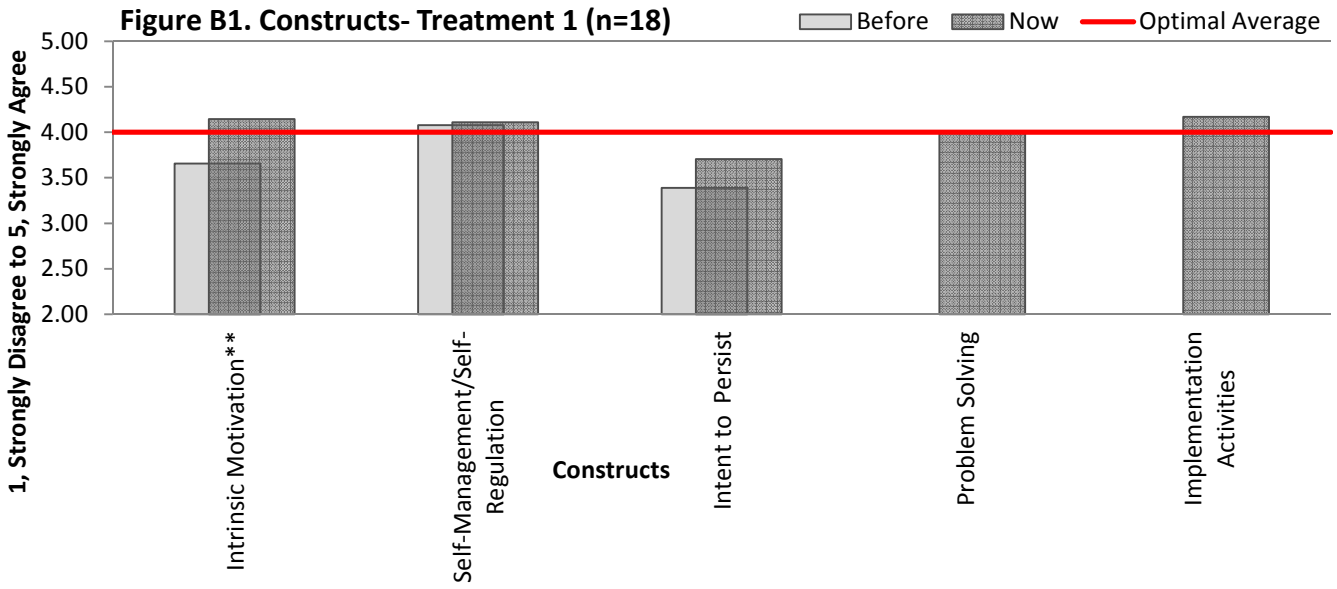
[7] Statistical power is the ability of a test to detect an effect, if the effect actually exists. Statistical power is contingent on an adequate sample size and the effect size (the salience of the treatment relative to the noise in measurement).

Figures B1 and B2. This suggests that both treatments are maximizing their impact on students' attitudes; however, additional attention may be needed in enhancing students' intentions to persist in a STEM-related field.

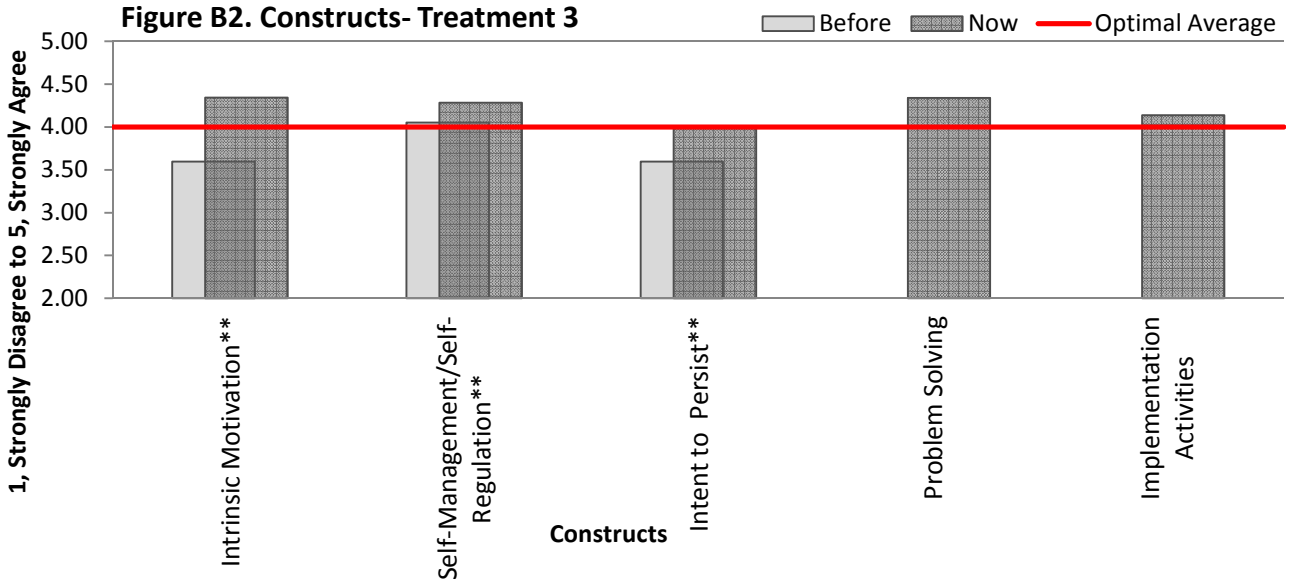
Table B1. Summary of Results by Constructs

Overall- Constructs							
		Treatment 1: Full Scientific Research Course			Treatment 3: Module/unit only- First year		
Constructs		n	Mean	Paired Samples t-test	n	Mean	Paired Samples t-test
Intrinsic Motivation	Before	18	3.65	p<0.001**	194	3.60	p<0.001**
	Now	18	4.14		194	4.34	
Self-Management/ Self-Regulation	Before	18	4.08	p=0.836	193	4.05	p<0.001**
	Now	18	4.11		192	4.28	
Intent to Persist	Before	18	3.39	p=0.160	191	3.59	p<0.001**
	Now	18	3.70		192	3.99	
Problem Solving	Now	18	3.99	N/A	192	4.34	N/A
Implementation Activities	Now	18	4.17	N/A	192	4.14	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.



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.

2. STEM for Life Carroll County

The STEM for Life program at Carroll County Schools consisted of students who have been in the program for one semester or less ("low dosage") and students who have been in the program for more than one semester ("high dosage").

Given the differences in dosage described above, the current analysis displays separate findings for each dosage group (low and high). Specifically, evaluators utilized students' responses to the following survey question to differentiate low and high dosage participants: *How long have you participated in this program?* The table below indicates that 104 students are in the low dosage group and 75 students are in the high dosage group. Students who responded "Don't Know" on the survey form were not included in either group.[8]

How long have you participated in this program?		n	Low Dosage	High Dosage
Dosage	0 semesters	21	21	--
	1 semester	80	80	--
	2 semesters	27	--	27
	3 semesters	18	--	18
	4 or more semesters	30	--	30
	Summer Only	3	3	--
	Don't Know	20	--	--
	Total	**199**	**104**	**75**

Table B2 summarizes students' responses per dosage condition. Among students in the low dosage group (e.g., participation for one semester or less), statistically significant increases were detected across the following constructs from before the program to now: *Intrinsic Motivation*, *Self-Management/Self-Regulation Skills* and *Intent to Persist*. The largest student gains observed were in the *Intrinsic Motivation* construct. Before the program, low dosage

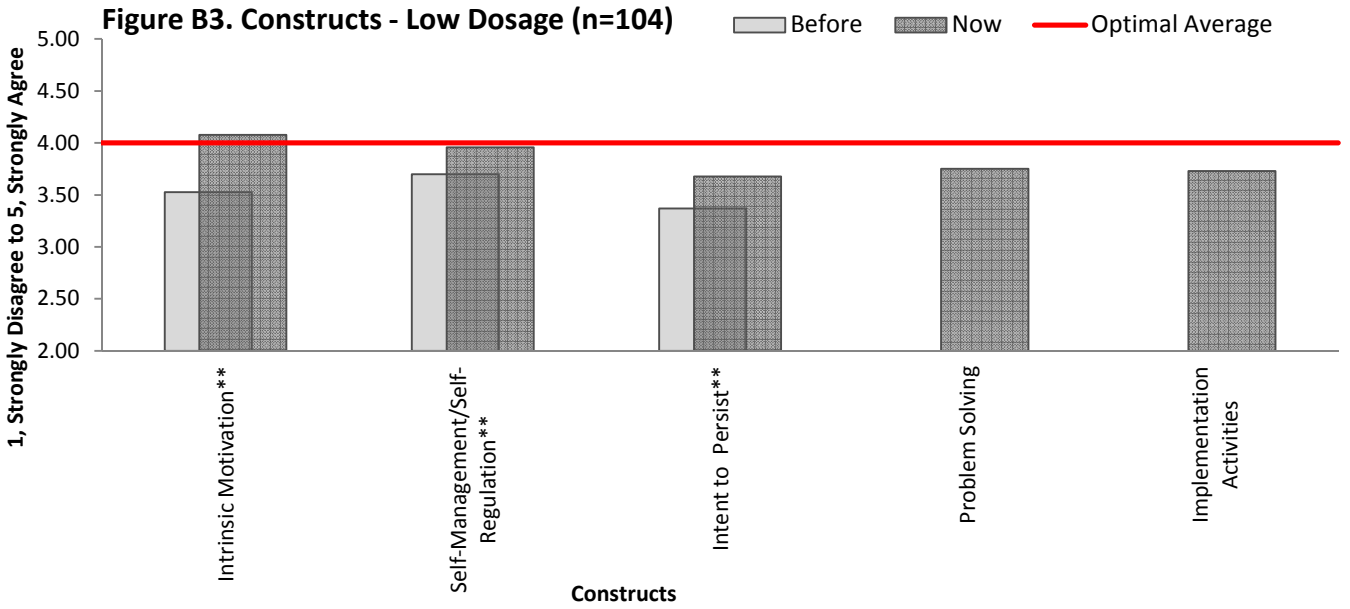
[8] Students who indicated "Don't Know" were excluded from the analysis. Because the survey asks students to reflect on their attitudes *before* the program and compare them to their attitudes *now*, students who are not clear on how long they participated in the program may have reported inaccurate data.

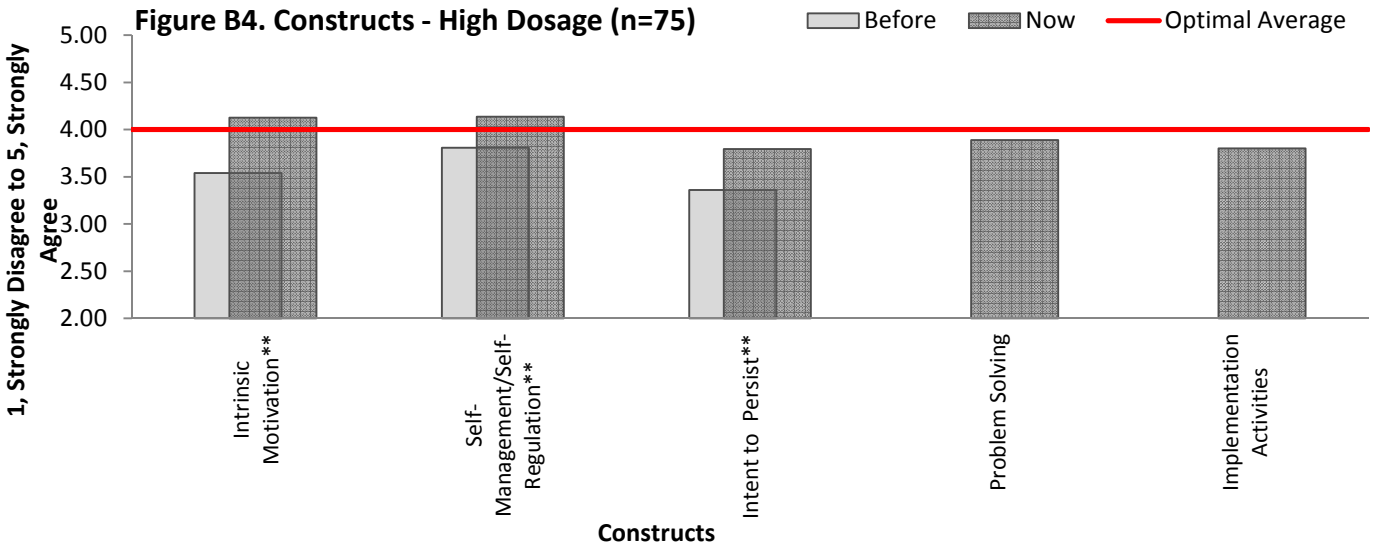
students rated their motivation to learn about STEM at a mean of 3.53 on a 5-point Likert scale (1, *Strongly Disagree* to 5, *Strongly Agree*); now, students indicate that they are intrinsically motivated to engage in STEM-related tasks and projects with a mean of 4.08 on a 5-point Likert scale. Among students in the high dosage group (e.g., participation for more than one semester), statistically significant gains were detected for the following constructs: *Intrinsic Motivation, Self-Management/Self-Regulation Skills,* and *Intent to Persist*. Like the students in the low dosage group, the largest gains for the high dosage students were in the *Intrinsic Motivation* construct. Before the program, high dosage students rated their motivation to learn about STEM at a mean of 3.54; now, after the program, students rate their motivation to learn STEM concepts a 4.13. Examining the "now" scores for the high and low dosage groups, it is evident that, across all constructs, students in the high dosage group show higher mean scores. This suggests that students who are part of the STEM for Life program for longer than one semester experience slightly more positive attitudes towards STEM than students who are part of the program for one semester or less. See Figures B3 and B4.

Table B2. Summary of Results by Constructs

Overall- Constructs							
		Low Dosage			High Dosage		
Constructs		**n**	**Mean**	**Paired Samples t-test**	**n**	**Mean**	**Paired Samples t-test**
Intrinsic Motivation	Before	103	3.53	p<0.001**	75	3.54	p<0.001**
	Now	100	4.08		72	4.13	
Self-Management/ Self-Regulation	Before	104	3.70	p<0.001**	75	3.81	p<0.001**
	Now	102	3.96		75	4.14	
Intent to Persist	Before	104	3.37	p<0.001**	75	3.36	p<0.001**
	Now	103	3.68		75	3.79	
Problem Solving	Now	104	3.75	N/A	75	3.89	N/A
Implementation Activities	Now	104	3.73	N/A	75	3.80	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.





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.