DeSIRE Consultation Report

Tyler Pollard

2024-07-02

Contents

Abstract

1 Introduction

1.1 Purpose

Engaging middle school students in rural North Carolina in STEM education presents a critical challenge. These students often face unique barriers, including limited access to resources, fewer educational opportunities, and less exposure to STEM based careers. This gap in engagement and access can suppress curiosity, limit career aspirations, and hinder the development of crucial skills needed in an advancing technological world. To overcome these challenges, innovative educational programs specifically designed to their needs are necessary. These programs aim to spark interest and provide practical learning experiences that enhance both cognitive outcomes like STEM knowledge and practices, and non-cognitive outcomes such as STEM interest, identity, and self-efficacy.

DeSIRE is a partnership between the College of Engineering and the Friday Institute for Educational Innovation at North Carolina State University (NCSU), the NC Mathematics and Science Education Network Pre-College Program (MSEN), the Edgecombe County Public Schools (ECPS) district, and local advanced manufacturing industry. The Friday Institute (FI) serves as the primary research partner for the DeSIRE project.

1.2 Program design

The aim of the DeSIRE project is to create community-based engineering design experiences for underserved middle school students (grades 6-8) from rural NC aimed to improve their cognitive (STEM content knowledge and career awareness) and non-cognitive (interest, self-efficacy and STEM identity) outcomes, and ultimately lead to their increased participation in STEM fields, particularly engineering.

The DeSIRE program involves middle school students from West Edgecombe and Phillips through mentoring and instructional activities centered around the engineering design process, specifically aligned with careers in the region's advanced manufacturing industries. By engaging students in project-based learning in school and providing enrichment activities during weekends and summer breaks, the initiative seeks to expand their understanding of STEM career paths, emphasizing opportunities in advanced manufacturing within their local communities.

DeSIRE will accomplish the overarching project goal by: 1) developing a 3-part Engineering Design elective course for grade levels 6-8, 2) incorporating a mentoring component whereby undergraduate engineering students from the Minority Engineering Program (MEP) at NC State and STEM professionals from industry serve as mentors to the middle school students during the course and 3) providing In-depth STEM Experiences where students engage in supplemental STEM enrichment activities outside of the classroom such as industry and university tours.

1.3 Research Questions

1. How and to what degree does the engineering design-focused program impact students' disciplinary- based knowledge and practices in STEM and career awareness?

2. How and to what degree does the engineering design-focused program impact students' STEM interest, STEM identity, and STEM self-efficacy?

3. How and to what degree does the engineering design-focused program impact teacher STEM content and pedagogical knowledge and awareness of STEM

2 Methods

2.1 Data

Table 1: Demographics of the middle school students in the S-STEM survey data

Variable	Category	Frequency	Percentage (%)
SchoolYear	2020-2021	5	6.1
	2022-2023	77	93.9
Semester	Fall	42	51.22
	Spring	40	48.78
YearSemester	2021Spring	5	6.1
	2022Fall	42	51.22
	2023Spring	35	42.68
School	Phillips Middle School	56	68.29
	West Edgecombe Middle School	26	31.71
Grade	6th	49	59.76
	7th	19	23.17
	8th	14	17.07
Gender	Female	39	47.56
	Male	39	47.56
	Other	4	4.88
Gender2	Male	39	47.56
	Not Male	43	52.44
Race	American Indian/Alaska Native	1	1.22
	Black/African American	46	56.1
	Hispanic/Latino	12	14.63
	Multiracial	5	6.1
	Other	4	4.88
	White/Caucasian	14	17.07
Race2	Black/African American	46	56.1
	Hispanic/Latino	12	14.63
	White/Caucasion	14	17.07
	Other	10	12.2
Total		**82**	**100**

2.2 Analysis of Variance

$$Y_{ijk} = \mu + \alpha_i + \beta_j + E_{ijk} \text{ where } E_{ijk} \overset{iid}{\sim} N(0, \sigma^2)$$
$$i = 1, 2 = a, \text{ for factor School}$$
$$j = 1, 2 = b, \text{ for factor Student}$$
$$k = 1, ..., n_{ij} \text{ for \# students per group}$$

Table 2: ANOVA Table

Source	Sum Sq	Df	F value	Pr(>F)
Student	1559.8	1	28.414	2.53e-07
School	8658.2	1	157.735	< 2.2e-16
Residuals	11472.2	209		

2.3 Ordinal Regression

$$Y_i = \begin{cases} \text{Not Proficient} & Y_i^* \leq \theta_1 \\ \text{Level 3} & \theta_1 < Y_i^* \leq \theta_2 \\ \text{Level 4} & \theta_2 < Y_i^* \leq \theta_3 \\ \text{Level 5} & \theta_3 < Y_i^* \end{cases}$$

for $i = 1, ..., N = 102$

for $j = 1, ..., 4$

$$Y_i^* \sim Normal(\mu_i, \sigma^2)$$

$$P(Y_i = \text{Not Proficient}|\mu_i, \sigma^2 = 1, \theta_j) = P(Y^* \leq \theta_1) = \Phi(\theta_1 - \mu_i)$$
$$P(Y_i = \text{Level 3}|\mu_i, \sigma^2 = 1, \theta_j) = P(\theta_1 < Y^* \leq \theta_2) = \Phi(\theta_2 - \mu_i) - \Phi(\theta_1 - \mu_i)$$
$$P(Y_i = \text{Level 4}|\mu_i, \sigma^2 = 1, \theta_j) = P(\theta_2 < Y^* \leq \theta_3) = \Phi(\theta_3 - \mu_i) - \Phi(\theta_2 - \mu_i)$$
$$P(Y_i = \text{Level 5}|\mu_i, \sigma^2 = 1, \theta_j) = P(\theta_3 < Y^*) = 1 - \Phi(\theta_3 - \mu_i)$$

$$\begin{aligned} \mu_i &= \beta_0 + \beta_1(\text{sexF}_i) + \beta_2(\text{ethnicH}_i) + \beta_3(\text{ethnicM}_i) + \beta_4(\text{ethnicW}_i) \\ &\quad + \beta_5(\text{dosage}_i) + \beta_6(\text{dosage}_i^2) + \beta_7(\text{Level\_5}_i) + \beta_8(\text{Level\_5}_i^2) \\ &\quad + \beta_9(\text{dosage}_i)(\text{Level\_5}_i) + \beta_{10}(\text{dosage}_i^2)(\text{Level\_5}_i) \\ &\quad + \beta_{11}(\text{dosage}_i)(\text{Level\_5}_i^2) + \beta_{12}(\text{dosage}_i^2)(\text{Level\_5}_i^2) \\ \beta_0 &= 0 \\ \sigma^2 &= 1 \\ \beta_1, .., \beta_{12} &\sim \text{Normal}(0, 10), \text{ uninformative priors} \\ \theta_1, ..., \theta_3 &\sim \text{Normal}(0, 10), \text{ uninformative priors} \end{aligned}$$

2.4 Logistical Regression Modeling

$$
\begin{aligned}
E(Y_i|\mathbf{X}_i) &= P(Y_i = 1) = \pi_i \\
\text{logit}(\pi_i) &= \log\left(\frac{\pi_i}{1-\pi_i}\right) = \log(\text{Odds}_i) \text{ for } i = 1, ..., n = 70, \text{ for number of students} \\
\log(\text{Odds}_i) &= \beta_0 + \beta_1 \text{SchoolPhillipsMiddleSchool}_i + \beta_2 \text{Grade7th}_i \\
&\quad + \beta_3 \text{Grade8th}_i + \beta_4 \text{Gender2NotMale}_i + \beta_5 \text{Race2Black\_AfricanAmerican}_i \\
&\quad + \beta_6 \text{Race2White\_Caucasian}_i + \beta_7 \text{Race2Hispanic\_Latino}_i
\end{aligned}
$$

2.5 Non-parametric tests of group distributions

2.6 Bayesian regression modeling

$$
\begin{aligned}
Y_i &\sim \text{Normal}(\theta_{t_i} + \mu_i, \sigma^2) \\
t &= 1, 2, 3 \quad \text{for the three S-STEM survey times} \\
i &= 1, ..., n \quad \text{for the number survey responses}
\end{aligned}
$$

$$
\begin{aligned}
\mu_i &= \beta_0 + \beta_1(\text{SchoolPhillipsMiddleSchool}_i) \\
&\quad + \beta_2(\text{Grade7th}_i) + \beta_3(\text{Grade8th}_i) + \beta_4(\text{Gender2NotMale}_i) \\
&\quad + \beta_5(\text{Race2Black\_AfricanAmerican}_i) + \beta_6(\text{Race2Hispanic\_Latino}_i) \\
&\quad + \beta_7(\text{Race2White\_Casucasian}_i)
\end{aligned}
$$

$$
\begin{aligned}
\beta_0 &\sim \text{Normal}(0, 10) \quad \text{for the intercept} \\
\beta_1, ..., \beta_7 &\sim \text{Normal}(0, 10) \quad \text{for fixed effects} \\
\theta_{t_i} &\sim \text{Normal}(0, \tau^2) \quad \text{Random effect for survey time } t \text{ of response } i \\
\sigma^2 &\sim \text{InvGamma}(0.1, 0.1) \\
\tau^2 &\sim \text{InvGamma}(0.1, 0.1)
\end{aligned}
$$

3 Results

3.1 Impact on disciplinary-based knowledge through EOG science test

3.2 STEM students' career awareness

3.2.1 Scientists

Expected Marginal Probability of Career Awareness for Scientists

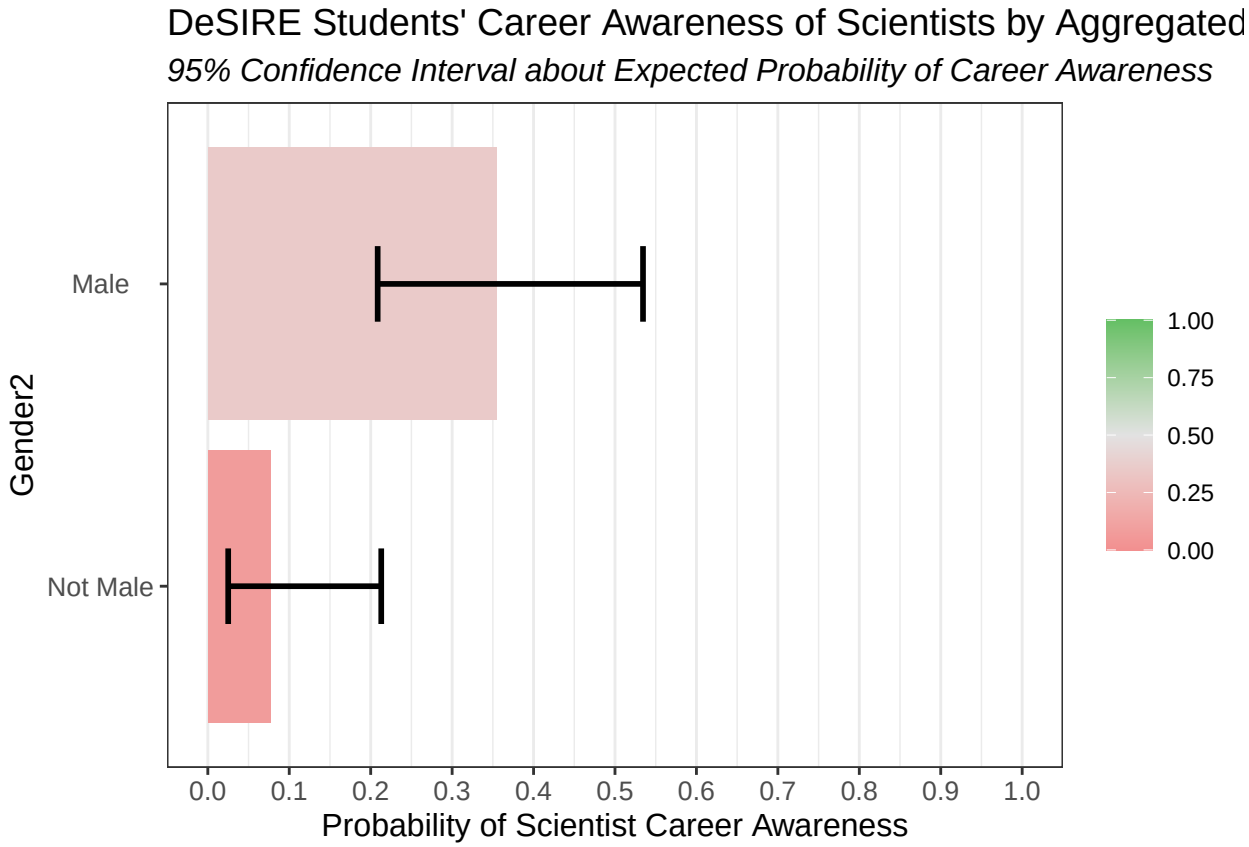
Variable	EMP[1]	95% CI[2]
Gender		
Male	0.36	(0.21, 0.53)
Not Male[3]	0.08	(0.03, 0.21)

[1] EMM = Expected Marginal Probability
[2] CI = Credible Interval
[3] Includes Females and Other

Table 3: Model Summary for Scientist Career Awareness

Characteristic	N	OR[1]	SE[1]	95% CI[1]	p-value[2]
Gender2	70				**0.003****
Male	31	1.00	—	—	
Not Male	39	0.15	0.709	0.03, 0.55	

[1]OR = Odds Ratio, SE = Standard Error, CI = Confidence Interval
[2]*p<0.05; **p<0.01; ***p<0.001


DeSIRE Students' Career Awareness of Scientists by Aggregated
95% Confidence Interval about Expected Probability of Career Awareness
Gender2
Male
Not Male
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Probability of Scientist Career Awareness
1.00
0.75
0.50
0.25
0.00

3.2.2 Engineers

3.2.3 Mathematicians

3.2.4 Technologists

3.3 Students' STEM interest

3.4 Students' STEM identity

Significant Kruskal-Wallis H Tests for All Interest Areas

Interest Area	Variable	χ^2	Df[1]	p-value[2]
Physics	Gender2	4.63	1	0.0313*
	Race2	13.43	3	0.0038**
Biology and Zoology	School	3.91	1	0.0481*
Veterinary Work	School	4.6	1	0.0319*
Mathematics	School	4.59	1	0.0321*
	Race2	8.25	3	0.0411*
Earth Science	Race2	8.84	3	0.0315*
Computer Science	Gender2	4.01	1	0.0451*
Medical Science	Gender2	4.05	1	0.0441*
Chemistry	Gender2	6.19	1	0.0128*
	Race2	8.34	3	0.0395*
Energy	Gender2	8.78	1	0.003**
	Race2	11.72	3	0.0084**
Engineering	Gender2	8.69	1	0.0032**

[1] Df = Degrees of Freedom

[2] $p < 0.05^*$; $p < 0.01^*$; $p < 0.001^*$

3.5 Students' STEM self-efficacy

3.5.1 Math Construct

Model Summary for Math Self-Efficacy

Variable	N	β	SD[1]	95% CI[2]
Intercept	82	3.08	0.1	(2.89, 3.27)

[1] SD = Standard Deviation

[2] CI = Credible Interval

3.5.2 Science Construct

Model Summary for Science Self-Efficacy

Variable	N	β	SD[1]	95% CI[2]
Intercept	82	3.91	0.16	(3.6, 4.21)
School				
West Edgecombe Middle School	26	–	–	–
Phillips Middle School	56	-0.39	0.17	(-0.72, -0.06)
Gender				
Male	39	–	–	–
Not Male	43	-0.34	0.16	(-0.64, -0.04)

[1] SD = Standard Deviation
[2] CI = Credible Interval

Expected Marginal Means for Science Self-Efficacy

Variable	EMM[1]	95% CI[2]
School		
West Edgecombe Middle School	3.74	(3.48, 4.03)
Phillips Middle School	3.35	(3.15, 3.54)
Gender		
Male	3.71	(3.48, 3.93)
Not Male	3.38	(3.16, 3.61)

[1] EMM = Expected Marginal Mean
[2] CI = Credible Interval

3.5.3 Engineering and Technology

Model Summary for Engineering and Technology Self-Efficacy

Variable	N	β	SD[1]	95% CI[2]
Intercept	76	4.37	0.3	(3.78, 4.97)
School				
West Edgecombe Middle School	23	–	–	–
Phillips Middle School	53	-0.81	0.25	(-1.3, -0.31)
Race				
Other	9	–	–	–
Black/African American	42	-0.15	0.3	(-0.74, 0.44)
White/Caucasian	14	-1.12	0.36	(-1.83, -0.41)
Hispanic/Latino	11	-0.39	0.37	(-1.12, 0.34)

[1] SD = Standard Deviation
[2] CI = Credible Interval

Expected Marginal Means for Engineering and Technology Self-Efficacy

Variable	EMM[1]	95% CI[2]
School		
West Edgecombe Middle School	3.95	(3.57, 4.33)
Phillips Middle School	3.14	(2.86, 3.43)
Race		
Other	3.96	(3.44, 4.49)
Black/African American	3.81	(3.52, 4.1)
White/Caucasian	2.84	(2.38, 3.31)
Hispanic/Latino	3.57	(3.05, 4.12)

[1] EMM = Expected Marginal Mean
[2] CI = Credible Interval

4 Conclusion

Appendix A: R Code to clean S-STEM data

```
# Load Libraries ---- Data Reading
library(data.table)  # Read csv in tibble format
library(readxl)  # Read xlsx

## Data Manipulation
library(stringr)  # Manipulate strings
library(plyr)  # Produce summary tables/data.frames

## Data Analysis
library(likert)
library(lme4)
library(lmtest)
library(car)
library(emmeans)
library(betareg)
library(caret)

## Bayesian Data Analysis
library(DescTools)
library(rstanarm)
library(brms)
library(posterior)
library(bayesplot)
library(BayesFactor)
```

```r
## Plotting
library(patchwork)

## Load this package last to reduce package conflictions
## with dplyr
library(tidyverse)

# %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
# S-STEM SURVEY
# ============================================================
# %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
# Read in Data ----

### Questions ----
SSTEMsurvey_questions <- read_excel("Data/S-STEM+-+DeSIRE_June+6,+2024_09.13.xlsx",
    n_max = 1)
SSTEMsurvey_questions <- SSTEMsurvey_questions |>
    unlist()
view(SSTEMsurvey_questions)
SSTEMsurvey_colnames <- names(SSTEMsurvey_questions)

### Survey Data ----
SSTEMsurvey_data_withQuestions <- read_excel("Data/S-STEM+-+DeSIRE_June+6,+2024_09.13.xl
    na = c("", NA), skip = 1)
SSTEMsurvey_data_withCodes <- SSTEMsurvey_data_withQuestions
colnames(SSTEMsurvey_data_withCodes) <- SSTEMsurvey_colnames

## Clean Data ----

### Select and Rename ----
SSTEMsurvey_data <- SSTEMsurvey_data_withCodes |>
    select(StartDate, Progress, Finished, str_which(SSTEMsurvey_colnames,
        pattern = "Q")) |>
    #### Socio-Demographics ----
rename(StartDate = StartDate, School = Q2, Grade = Q3, TeacherID = Q37,
    TeacherName = Q37_4_TEXT, SemesterYear = Q36, FirstName = Q15,
    MiddleName = Q16, LastName = Q17, BirthDate = Q18, StudentID = Q20,
    Gender = Q21, GenderOther = Q21_6_TEXT, Race = Q22, RaceOther = Q22_8_TEXT)

#### Self-Efficacy ----

## Math, Science, and Engineering and Tech Constructs
colnames(SSTEMsurvey_data) <- str_replace(colnames(SSTEMsurvey_data),
    pattern = "Q6_", replacement = "Math_Q")
```

```r
colnames(SSTEMsurvey_data) <- str_replace(colnames(SSTEMsurvey_data),
    pattern = "Q23_", replacement = "Science_Q")
colnames(SSTEMsurvey_data) <- str_replace(colnames(SSTEMsurvey_data),
    pattern = "Q24_", replacement = "EngTech_Q")

#### Identity ----

## 21st Century Learning Construct
colnames(SSTEMsurvey_data) <- str_replace(colnames(SSTEMsurvey_data),
    pattern = "Q26_", replacement = "Learning_Q")

#### Interest ----

## Your Future Items
interestAreas <- c("Physics", "Environmental Work", "Biology and Zoology",
    "Veterinary Work", "Mathematics", "Medicine", "Earth Science",
    "Computer Science", "Medical Science", "Chemistry", "Energy",
    "Engineering")

colnames(SSTEMsurvey_data) <- str_replace(colnames(SSTEMsurvey_data),
    pattern = "Q28_", replacement = "Interest_Q")

#### Career Awareness ----

## More about you Items
careerAwarenessAreas <- c("Scientists", "Engineers", "Mathematicians",
    "Technologists")

colnames(SSTEMsurvey_data) <- str_replace(colnames(SSTEMsurvey_data),
    pattern = "Q35_", replacement = "Awareness_Q")

#### Other Questions ----

##### Expectations ----
expectationAreas <- c("English", "Math", "Science")

colnames(SSTEMsurvey_data) <- str_replace(colnames(SSTEMsurvey_data),
    pattern = "Q29_", replacement = "Expectations_Q")

##### Take Advanced Classes ----
advancedClassesAreas <- c("Mathematics", "Science")

colnames(SSTEMsurvey_data) <- str_replace(colnames(SSTEMsurvey_data),
    pattern = "Q31_", replacement = "Classes_Q")
```

```r
##### College ----
colnames(SSTEMsurvey_data) <- str_replace(colnames(SSTEMsurvey_data),
    pattern = "Q32", replacement = "PlanCollege")
colnames(SSTEMsurvey_data) <- str_replace(colnames(SSTEMsurvey_data),
    pattern = "Q33", replacement = "InterestCollege")
colnames(SSTEMsurvey_data) <- str_replace(colnames(SSTEMsurvey_data),
    pattern = "Q34", replacement = "FirstCollege")

#### SAVE CHECKPOINT ----
save(SSTEMsurvey_data, file = "Data/S-STEM Survey Data Renamed.RData")

### Recode ----
load(file = "Data/S-STEM Survey Data Renamed.RData")

#### Date ----
SSTEMsurvey_data2 <- SSTEMsurvey_data |>
    mutate(Date = as_date(StartDate), Year = year(Date), Semester = ifelse(semester(Dat
        1, "Spring", "Fall"), SchoolYear = ifelse(Semester ==
        "Fall", paste0(Year, "-", Year + 1), paste0(Year - 1,
        "-", Year)), YearSemester = paste0(Year, Semester)) |>
    select(StartDate, Date, Year, SchoolYear, Semester, YearSemester,
        everything())

#### Factors of Interest ----
SSTEMsurvey_data3 <- SSTEMsurvey_data2 |>
    mutate(School = factor(School, labels = c("West Edgecombe Middle School",
        "Phillips Middle School")), Grade = factor(Grade, labels = c("6th",
        "7th", "8th")), Gender = factor(Gender, labels = c("Male",
        "Female", "Other")), Race = factor(Race, levels = c(1,
        2, 3, 4, 5, 6, 7, 8), labels = c("American Indian_Alaska Native",
        "Asian", "Black_African American", "Native Hawaiian_Other Pacific Islander",
        "White_Caucasian", "Hispanic_Latino", "Multiracial",
        "Other")), Race2 = factor(Race, levels = c("American Indian_Alaska Native",
        "Asian", "Black_African American", "Native Hawaiian_Other Pacific Islander",
        "White_Caucasian", "Hispanic_Latino", "Multiracial",
        "Other"), labels = c("Other", "Other", "Black_African American",
        "Other", "White_Caucasian", "Hispanic_Latino", "Other",
        "Other"))) |>
    select(1:21, Race2, everything())

SSTEMsurvey_data <- SSTEMsurvey_data3

### SAVE DATA ----
save(SSTEMsurvey_data, SSTEMsurvey_colnames, SSTEMsurvey_questions,
```

```
    file = "Data/Cleaned S-STEM Survey Data.RData")
```