The Future of Engineering Education
2024 Annual Conference & Exposition
Oregon Convention Center
Portland, OR . June 23 - 26, 2024
Paper ID #41290

STEM Interest as an Indicator of Elementary and Middle School Aged Youth's Decision to Participate in Out-of-School Informal STEM Education

Dr. Turhan K. Carroll, University of Georgia

Dr. Turhan Carroll is an Assistant Professor of Workforce Education at the University of Georgia. His research interests reside in two fields: engineering education and STEM education. His research agenda focuses on understanding the impacts of informal STEM education; supporting engineering education pathways for racially minoritized and economically disadvantaged pre-college students; and elementary STEM education.

Dr. Carroll earned his PhD in Engineering Education with an interdisciplinary specialization in Quantitative Research Evaluation and Methodology from The Ohio State University. Prior to joining the College of Education, he was a postdoctoral researcher in the College of Education and Human Ecology at The Ohio State University. Prior to his graduate work he worked as an engineer at the Air Force Research Laboratory in Dayton, OH.

Dr. Jessica R Hoehn, University of Colorado Boulder

Dr. Jessica R. Hoehn is a postdoctoral researcher at University of Colorado Boulder. She received her PhD in Physics Education Research from CU, studying ontological, epistemological, and social aspects of student reasoning in quantum mechanics. Dr. Hoehn

Prof. Noah D Finkelstein, University of Colorado Boulder

Noah Finkelstein is a Professor and Vice Chair in the department of Physics at the University of Colorado Boulder. He conducts research is in physics education, specifically studying the conditions that support students' identities, engagement and outcomes in physics - developing models of context. In parallel, he conducts research on how educational transformations get taken up, spread, and sustained. He is a PI in the Physics Education Research (PER) group and was founding co-director of CU's Center for STEM Learning. He co-directs the national Network of STEM Education Centers, is building the STEM DBER-Alliance, and coalitions advancing undergraduate education transformation. He is involved in education policy serving on many national boards, sits on a National Academies' STEM education roundtable, is a Trustee of the Higher Learning Commission, is a Fellow of both the American Physical Society and the American Association for the Advancement of Science, and a Presidential Teaching Scholar and the inaugural Timmerhaus Teaching Ambassador for the University of Colorado system.

STEM Interest as an Indicator of Elementary and Middle School Aged Youth's Decision to Participate in Out-of-School Informal STEM Education

Abstract

Youth typically decide whether to pursue science, technology, engineering, and mathematics (STEM) careers as early as middle school, suggesting that nurturing STEM interest in elementary and middle (primary) school is a key factor in attracting youth to engineering. Goals of racial equity and attracting youth into engineering have birthed the proliferation of many informal STEM education (ISE) programs (e.g., out-of-school programs, summer camps, etc.). Though research suggests that ISE increases participants' STEM interest, it is unclear whether ISE is successful in sparking STEM interest in previously uninterested youth. This gap exists partly because little is known about the initial STEM interest of ISE participants.

Using a survey research design, we addressed this gap by studying initial STEM interest among 336 primary school youth from the mountain west region of the United States: 44 of whom participated in the Partnerships for Informal Science Education in the Community (PISEC) program, and 292 who did not. The research questions guiding this study were:

> RQ1. To what extent do youth who did or did not participate in the PISEC program differ in their initial STEM interest?
>
> RQ2. Controlling for STEM identity, performance, recognition and future-self, to what degree is initial STEM interest predictive of youths' decision to participate in the PISEC program?

This research contributes to pre-college engineering education scholarship by deepening our understanding of youth who participate in ISE and illuminating ways to better attract those uninterested in STEM. By providing insight into the baseline interest of ISE participants, our research furthers the field's understanding of the long-term outcomes of ISE.
We used a STEM Identity survey containing Interest, Performance, Recognition, and Future-Self constructs. For RQ1, we conducted a two-tailed independent samples t-test. We found a moderately large difference in STEM interest between ISE participants and non- participants in this study $(t(131.805) = -8.764, p < .001, d = .63)$. Using logistic regression for RQ2, we found that youth with high initial STEM interest are around 5 times more likely to choose to participate in PISEC than youth with lower STEM interest $(OR_{Interest} = 5.235)$.

These findings provide initial evidence that youth attracted to ISE have high initial STEM interest. Understanding this allows ISE stakeholders to develop strategies to both attract youth who are uninterested in STEM and support those who already have interest. In this way, our study begins to elucidate a STEM identity profile for youth who currently elect to participate in

this program and how recruiting strategies could be tailored in ways that strengthen ISE's ability to broaden participation in STEM. Moreover, reaching students with higher STEM interest is beneficial because it can prevent their exclusion from STEM. Attracting these youth to ISE can foster positive STEM identity formation among them. We report the details of the study, methods, analyses and findings. We then discuss implications of these findings for pre-college

Introduction

The number of jobs in the US requiring training in science and engineering is on the rise, yet the number of students receiving training in these fields is declining at an alarming rate [1], [2]. It has been long believed that too few undergraduates are recruited and retained in STEM programs to meet the nations need. It was historically thought that the first two years of college are the most critical to the retention and recruitment of STEM majors [3]. However, in recent years, a consensus among scholars has emerged that efforts to recruit and retain students into STEM fields should begin as early as primary (e.g. elementary or middle) school. In fact, prior work has found that students tend to decide whether or not to pursue STEM careers by middle school [4]. This means that it is important to initiate STEM interest among youth during their primary grades.

One promising vehicle for initiating STEM interest among youth is informal STEM education (ISE). Informal STEM education is STEM education that occurs in informal learning environments. In STEM education research literature, informal learning environments are defined as any STEM learning environment that exists outside of a traditional classroom [5], [6]. In their 2009 policy report, National Academy of Sciences (NAS) Committee on *Learning Science in Informal Environments* Bell et al. [5] delineated three types of informal learning environments relevant to K12 learners: everyday environments, designed environments, and out-of-school environments. Table 1 describes some examples and key characteristics of each of these environments.

Table 1. NAS categories for informal learning environments.

Venue	**Examples**	**Important Characteristics**
Everyday Environments	Family discussion, hobbies, technology use	Interaction coordinated by the learner; learning varies across and within cultures; learning and assessment are not easily distinguished.
Designed Environments	Museums, zoos, science centers, libraries	Artifacts, media, and signage guide the learner's experience; learner's interaction with the environment is often determined by the individual; learner engagement is short-term and sporadic in the setting; learning tends to take place in peer, family, or mentor interactions.
Out-of-School Environments	After-School Programs, STEM Summer camps, science center programs.	Often consists of a designed curriculum; focus may be on content knowledge or solving applied problems; learning typically guided and/or monitored by a trained facilitator; assessment sometime used but are not used to judge individual's performance or learning by an institutional standard.

Bell et al. [5] delineated six learning outcomes that result from participation in informal science environments. Two that they argued had less overlap with learning outcomes of formal schooling and thus were particularly relevant to informal environments were: the ability of participants to experience excitement, interest, and motivation to learn about the physical world; and the ability of participants to develop an identity as someone who knows about, uses, and can contribute to science. In the years since the release of this report, a substantial body of research has emerged providing evidence that participation in informal STEM education is positively associated with the decision of youth to persist in STEM and pursue STEM careers [4], [7], [8]. Researchers have further found that early exposure to STEM through informal STEM education has a positive effect on STEM identity formation among participants [9], [10], [11]. Much of this work on STEM identity formation, for example Dou et al. [9] also finds that STEM interest is a key component of STEM identity formation. This suggests that STEM interest and STEM identity are interrelated.

The literature clearly lays out the positive impact that informal STEM education has on STEM interest and STEM identity formation among youth. What is less clear from the literature is whether or not ISE is successful is successful in attracting previously uninterested youth. In other words, are ISE programs initiating interest in STEM, nurturing an already strong interest in STEM, or both? If ISE researchers and stakeholders hope to use these programs to help attract and train the future STEM workforce and broaden participation in STEM, it is vital that work to attract youth who are previously uninterested in STEM. One cause of this gap in literature is the fact that little is known about the initial STEM interest of ISE participants. In order to address this gap, we have performed a study in which we investigated the baseline STEM interest of primary-aged participants who participated in an out-of-school ISE program called PISEC and compared them with a control group of primary-aged youth who did not participate in the program. The research questions guiding our study were:

> RQ1. To what extent do youth who did or did not participate in the PISEC program differ in their initial STEM interest?
>
> RQ2. Controlling for STEM identity, performance, recognition and future-self, to what degree is initial STEM interest predictive of youths' decision to participate in the PISEC program?

This work contributes to scholarship on pre-college engineering education by providing a clearer understanding of interest level of youth who are currently participating to ISE giving us insight into who we are attracting and how we might attract youth with varied levels of STEM interest. By developing this understanding of baseline STEM interest, our work sets the stage for understanding of long-term impact of ISE on STEM interest.

Theoretical Framework

This study is grounded in two theoretical constructs: STEM interest and STEM identity. STEM interest can be thought of as someone's desire to continually engage with STEM activities and topics. STEM identity refers to a person's recognition, by self and others, as a STEM person [12]. In this way we might view STEM interest a key contributor in initiating positive STEM identity formation. We begin this section by discussing our framing of interest. We then discuss our framing of STEM identity including the key dimensions of our framework. This section concludes with a discussion of the relationship between STEM identity and STEM interest.

Interest

Following the work of Maltese et al (2014), we frame interest using Hidi and Renninger's four-phase model of interest development [14]. According to the four-phase model of interest development, phase 1 is Triggered Situational Interest; phase 2 is Maintained Situational Interest; phase 3 is Emerging Individual Interest; and phase 4 is Well-Developed Individual Interest. Since this paper is concerned primarily with baseline interest of individuals participating in an ISE program, we will focus the rest of our discussion on phase 1 of the four-phase model of interest. According to the theory, phase 1 is a psychological state of interest resulting from short-term changes in affective and cognitive processing. In this model the Triggered Situational Phase has the following four defining characteristics [14]:

1. it can be sparked by environment features such as personal relevance;
2. triggered situational interest is typically externally motivated;

3. learning environments that include group work, puzzles, or computers tend to trigger situational interest; and
4. triggered situational interest may be a precursor to reengaging in specific content (such as STEM content) over time as one advances through the other phases of interest.

Hidi and Renninger's four-phase model of interest is appropriate for the current study because as mentioned above we are interested in understanding the relationship between initial STEM interest and ISE participation with the aim of understanding the long-term impact of ISE on the development of participants' STEM interest. Though this current study represents a cross-section at the phase 1 of Hidi and Renninger's model, this phase represents an important phase for introducing youth to STEM helping them begin to see themselves as people who can do STEM. Moreover, the first phase of interest may be the point in which youth begin to form a positive STEM identity.

STEM Identity

There are multiple ways in which STEM identity is operationalized in informal STEM education research literature, with no consensus on one way to frame identity formation in STEM. For instance, Dou et. al. [9] frame STEM identity in terms of STEM interest and STEM recognition. Hughes et. al. (2013) who operationalize STEM identity as a construct containing of three dimensions: interest in STEM; self-concept as it relates to STEM subjects; and role models' impacts on students' perceptions of STEM professionals. For this work we operationalized STEM identity using Hazari et. al.'s (2010) physics identity framework. In our adaptation of Hazari et al.'s (2010) framework, STEM identity is made up of four dimensions:

1. Recognition: recognition by others as being good at STEM;
2. Interest: Desire/curiosity to think about and understand STEM;
3. Performance: belief in ability to perform required STEM tasks; and
4. Competence: belief in ability to understand STEM content.

For the remainder of this paper, when the term STEM identity is used, this is the definition we are using. Further, when the terms interest (or STEM interest) are used throughout the remainder of this paper, it is defined as stated in (2) above. STEM identity has been found to be an important factor in persistence of students toward STEM degrees [17]. There is an extensive body of literature on the topic of identity formation in informal STEM education [18], [19], [20], [21].

Dou et. al. (2019) studied the link between childhood Informal STEM learning experiences are predictive of STEM identity in college. They found that after controlling for home environment, gender, and other demographic factors, talking with friends and family about science, and consuming popular science media had significant effects on STEM identity among college students. Similarly, Riedinger and McGinnis (2017) found that students in a youth camp were authored identities as learners of science through learning conversations and performance. These findings suggest that that the life-wide nature of informal STEM learning, proposed by the National Academy of Sciences [5] contributes to STEM identity formation.

Summary

To summarize, this study uses Hidi and Renninger's four-phase model of interest and Hazari et al.'s physics identity framework as a lens to understand STEM interest within an informal STEM setting. While we utilized all four tenets of Hazari et al.'s physics identity work as we operationalized the theoretical framework for our study, we conceived of interest using phase 1 of the four-phase interest model. This was done because for this study we were primarily concerned with initial interest, however we used this framework because the current research is part of a larger study that will study the development of STEM interest and STEM identity over time. According to the four-phase theory of interest development, we can reasonably expect that youth who continually reengage in PISEC programming over time will progress through all four phases of interest development.

We operationalized STEM identity using a generalized, modified, shortened version of Hazari et. al.'s (2010) physics identity instrument. Our instrument is called the Primary STEM identity instrument and contains 5 constructs: Interest, Recognition, Performance/Competence, Future Selves, and Sense of Belonging in PISEC (this construct is not reported in this study because it did not apply to our control group). By using this instrument, we were able to assess students baseline STEM interest, STEM identity, and sense of belonging in the program in one brief instrument.

Literature Review

Informal STEM Education and STEM Interest

Informal STEM education has been demonstrated to have multiple positive impacts on youth. Informal STEM education is known to increase students' interest in pursuing STEM careers. In studying the intersections between family education, informal science experiences, and initial interest in science Dabney et. al. (2016) found that parental education level was predictive of early interest in science. Further, they found that the following forms of informs of science education were associated with early interest in science by elementary school: diversions, hobbies, and parental encouragement. Kong et. al. [24] found that participation in summer camps is positively associated with career interest in STEM fields. Bicer and Lee (2023) found that a two-week long STEM summer camp significantly increased participants' interest in pursuing math and science majors and careers and concluded that the camp has potential to increase students' interest in certain STEM fields. In addition to the work cited here, other scholars have reported on the link between informal STEM participation and increased STEM interest [25], [26], [27], [28].

Informal STEM Education, STEM Identity, and Broadening Participation in STEM

Given that the vast majority of the participants in PISEC are Hispanic/Latino youth, we now focus our literature review to research regarding ISE's impact on STEM identity formation for racially minoritized groups. King and Pringle [11] found that participation of Black girls in a community based informal STEM program reported developing interest in STEM learning. Young et al. (2019) used the High School Longitudinal Study (HSLS09/12) to examine the influence of informal STEM activities on math identity constructs for African American students. They found that Informal STEM participation showed a positive effect on students' math identity. Carlone et. al. (2015) performed an ethnographic study of minority youth participating in an ecology enrichment program for diverse youth. The study found, among other things, that various cultural norms, tools, and practices enabled the youth's identity development. This work is important because is illustrates the potential for ISE to truly broaden participation in STEM by attracting and preparing racially marginalized youth, while allowing them the space to form STEM identities. Henderson et al. [10] argued that there are three key mechanisms for engineering identity formation among Black boys: (1) practicing STEM (engaging in STEM-related activities); (2) exposure to STEM role models, and; (3) access to STEM mentors. Though research interleaving race and STEM identity in informal environments in relatively new, there has long been scholarship on STEM identity for racially minoritized groups [19], [30], [31], [32], [33], [34]. This prior work is vital in that it provides a framework for the elements necessary for positive identity formation for racial minorities as PISEC works to design programming that fosters positive STEM identity formation.

Methods

Program Context

Partnerships for Informal Science Education in the Community (PISEC) is a program run by the University of Colorado Boulder that connects university volunteers (undergraduate students, graduate students, and postdocs) with local K-12 students around hands-on, inquiry-based STEM activities. PISEC partners with schools and community organizations to run afterschool clubs for youth in grades 3-8, for which university volunteers travel to the partner site (i.e., school or community center) for one hour per week for 10 weeks each semester. PISEC works primarily with marginalized youth and aims to create opportunities for transformative and empowering experiences with science, facilitating pathways into STEM, and thereby helping to diversify STEM disciplines. Specific program goals include cultivating and maintaining youth interest in STEM and supporting youth development of a STEM identity and sense of belonging in STEM.

The PISEC afterschool program blends play and science learning and provides ample opportunities for youth participants to exercise agency over their scientific learning. Each semester, a given site runs one curriculum, which consists of a suite of 25-35 hands-on activities all centered around a physics topic (mechanics, electricity and magnetism, optics, thermodynamics, or astronomy). Youth participants work in groups with one PISEC mentor and each week they decide which experiments to do, how to do them, and in many cases, they design their own experiments. Throughout the semester, students document their experiments in a science notebook and also through 'vlog-style' video recordings, or live action or stop action motion movies. The semester then culminates with a field trip to PISEC. PISEC also hires experienced volunteers as Site Leaders; each site is led by a different site leader.

Data in this study come from two semesters of the PISEC afterschool program at four sites. Site A is K-8 school that enrolled 21 unique students (grades 3-6) in PISEC across the two semesters. The student body at Site A is 83.4% Hispanic/Latino, 12.5% white, and 80% economically disadvantaged. Site B is a middle school that enrolled 26 unique students (grades 6-8) in PISEC across the two semesters. The student body at Site B is 52.5% Hispanic/Latino, 40.5% white, and

53% economically disadvantaged. Site C is a middle school that only ran PISEC in the second semester; they enrolled 10 students (grade 7). The student body at Site C is 63.9% Hispanic/Latino, 31.2% white, and 76% economically disadvantaged. Due to current school district policies, we are unable to collect demographic information from students and families. We report the demographics of the schools and are confident that the participants in our dataset are representative of these. At Sites A, B, and C, participation in the PISEC program is entirely voluntary – students and families have the opportunity to sign up for afterschool programs each semester and PISEC is one of the options. Site D is a community organization that provides mentorship and robust programming for cohorts of students from first grade through post-secondary. PISEC ran an afterschool program both semesters with one cohort of 26 students (grades 3-4). The student population at Site D is 91% persons of color (84% are Latino/Latina/Latinx) and 100% from under-resourced communities; 78% of students in the organization are or go on to be first-generation college students. At Site D, students and their families have opted into the overall programming of the organization, which takes place every day after school. As part of that, they participate in the PISEC program once a week.

Data Collection

We collected survey data from PISEC participants at all four sites, as well as students from the same grades at Sites A and B who did not participate in PISEC. All PISEC participants took the survey during the first and last program sessions each semester. Before beginning any activities, a PISEC Site Leader administered hard copy surveys asking students to spend ~5 minutes sharing their thoughts about STEM. The survey was available in English and Spanish, and students were told they could complete it in whichever language they preferred. At each site, we had English- and Spanish-speaking PISEC volunteers who could help students read and respond to the survey items if needed. Parental consent forms for the research study were sent home with students; those who returned the form and whose parent/guardian opted to participate in the research study are included in our dataset. Across the four sites and two semesters, 44 students opted into the study and are included in the dataset analyzed here.

In order to collect data from the control group (i.e., students at the same schools in the same grades who were not participating in PISEC), we partnered with teachers who administered the survey in their classes and/or asked their colleagues to do so. In instances where PISEC participants were in the classes where the survey was given, the teachers removed them from the control dataset. For the control group, the pre-survey was given in the fall semester around the same time as the pre-survey in the PISEC program and the post-survey was given at the end of the spring semester. At Site A we received survey responses from 115 students in grades 4-6 and at Site B we received responses from 177 students in grades 6-8 (essentially the entire school as the science teachers for each grade administered the survey).

Data Analysis

Data analysis was performed using SPSS version 29.0.1. The variables used in this study, along with the coding scheme for each variable, is shown in Table 2.

Table 2: Coding Scheme for the Variables Used in this Study.

Variable	*Description*
PISEC	DV: 1=yes students participated in the PISEC program, 0= did not participate in the PISEC program
STEMID	1=no; 2=not sure; 3=yes. This is a one self-reported item about the respondent's perceived sense of STEM identity.
Interest	1=no; 2=not sure; 3=yes. This is a three item self-report construct about the respondent's STEM interest.
Recognition	1=no; 2=not sure; 3=yes. This is a two item self-report construct about the respondent's sense that authority figures see him or her as a STEM person.
Performance	1=no; 2=not sure; 3=yes. This is a three item self-report construct about the respondent's self-reported sense of self-efficacy toward STEM subjects.
Future Self	1=no; 2=not sure; 3=yes. This is a two item self-report construct about the respondent's self-reported desire to pursue a STEM career.
STEMID Score	Total score on STEMID Primary Survey (Average of all responses).

Since our first research question was concerned with a comparison between two independent groups, we used an independent samples t-test to answer it. For our second research question, the dependent variable (PISEC) has a dichotomous outcome, so we used logistic regression for our research question [35]. Logistic regression has the four major assumptions. First, the DV is a dichotomous variable. This is ensured by the selection of our DV. The second assumption is a linear relationship between the logit of the outcome and each continuous predictor. We used the Box-Tidwell test to check this assumption for the continuous variables [35], [36] and found all to be linear in the logit. The third assumption is that there are no extreme values or outliers. We used residual diagnostics to test this assumption [36] and found no extreme outliers. Finally, we assume that there is no multicollinearity among the predictors. We used linear regression to assess this [35] and below we discuss the results of our check of multicollinearity and how this impacted our final model.

Results

We conducted a study to investigate the extent to which STEM interest was predictive of decision to participate in the PISEC program among students attending schools in the rural mountain west region of the United States. Recall that two research questions guiding this study were:

> RQ1. To what extent do youth who did or did not participate in the PISEC program differ in their initial STEM interest?

RQ2. Controlling for STEM identity, performance, recognition and future-self, to what degree is initial STEM interest predictive of youths' decision to participate in the PISEC program?

We first produced frequency distributions for all our variables, also testing for normality. In order to test for normality, we overlaid a normal curve on each histogram, and we also included calculations of the skewness and kurtosis in the descriptive statistics of each variable for the study participants who did and did not participate in PISEC. The skewness and kurtosis for each variable are in Table 3 for non-PISEC participants and Table 3 for PISEC participants.

Table 3: Skewness and Kurtosis for the non-PISEC participants.

Predictor	n	Skewness	Kurtosis
STEMID	292	-.190	-1.160
Interest	292	-.842	-.522
Recognition	292	-.037	-.300
Performance	292	-.810	-.052
Future Self	292	-.219	-.856
STEMID_Score	292	-.668	-.454

Table 4: Skewness and Kurtosis for the PISEC participants.

Predictor	n	Skewness	Kurtosis
STEMID	44	-1.628	1.649
Interest	44	-3.306	10.741
Recognition	44	-.750	.202
Performance	44	-1.293	.426
Future Self	44	-1.017	.306
STEMID_Score	44	-2.439	7.983

According to the results shown in Table 3, our data for the non-PISEC participants is normally distributed (meaning that the skewness and kurtosis are both between -2.0 and 2.0). Table 4, however, suggests that the PISEC participants' data has a skewness and kurtosis that are outside of acceptable ranges for both the Interest construct and the total STEMID_Score. This is our first evidence that PISEC participants may skew toward higher levels of initial interest in STEM.

Next, we conducted t-tests for each independent variable to investigate differences between PISEC participants and non-PISEC participants. The t-tests results for the six variables are shown in the table below. The results in Table 5 suggest that all 6 variables showed statistically significant differences between students who participated in PISEC and students who did not. The results are in Table 5, and the results for our primary variable of concern (Interest) will be discussed further in the discussion section below.

Table 5: Independent Samples t-Tests for All Independent Variables

Predictor	t	df	p-value
STEMID	-4.624	334	<.001
Interest	-8.764	131.805 (Unequal Variances)	<.001
Recognition	-3.747	334	<.001
Performance	-6.705	86.517 (Unequal Variances)	<.001
Future Self	-3.927	334	<.001
STEMID_Score	-7.690	78.643 (Unequal Variances)	<.001

We conducted a two-tailed independent samples t-test to determine whether there were significant differences in self-reported STEM interest between PISEC participants and non-PISEC participants. Using Levene's test for equality of variances we determined that the variance in STEM interest between PISEC and non-PISEC participants was not equal ($F=57.077, p<.001$). Under this assumption, we found a medium to large statistically significant difference in STEM interest between primary PISEC participants ($M = 2.9$, $SD = .282$), and primary non-PISEC participants ($M = 2.4, SD = .665$), $t(131.805) = -8.764, p < .001, d = .63$. We conclude that there is sufficient evidence to support the notion that there are differences in STEM interest between students who participate in PISEC and students who are not.

Table 6 and Table 7 contain descriptive statistics for the non-PISEC participant group and the PISEC participant group, respectively, for each of the six predictors. It includes sample size (n), mean (M) and standard deviation from the mean (SD) for each predictor.

Table 6: Descriptive Statistics for the non-PISEC participants.

Predictor	n	M	SD
STEMID	292	2.120	.742
Interest	292	2.404	.665
Recognition	292	1.979	.534
Performance	292	2.397	.550
Future Self	292	2.068	.633
STEMID_Score	292	2.268	.506

Table 7: Descriptive Statistics for the non-PISEC participants.

Predictor	n	M	SD
STEMID	44	2.660	.608
Interest	44	2.909	.282
Recognition	44	2.307	.583
Performance	44	2.788	.322
Future Self	44	2.466	.575
STEMID_Score	44	2.708	.325

We used two methods to investigate the degree of multicollinearity for our predictor variables. We calculated the correlations between each predictor variable as shown in Table 8 and we performed ordinary least squares regression in order to calculation the variance inflation factors (VIFs) of each predictor (Table 9).

Table 8: Correlation Table for the Six Predictor Variables

Variable	n	M	SD	1	2	3	4	5	6
1. STEMID	336	2.190	.748	-					
2. Interest	336	2.470	.651	.640**	-				
3. Recognition	336	2.022	.551	.472**	.411**	-			
4. Performance	336	2.448	.541	.544**	.664**	.454**	-		
5. Future Self	336	2.121	.639	.527**	.613**	.423**	.515**	-	
6. STEMID_Score	336	2.325	.508	.762**	.909**	.589**	.800**	.785**	-

**. Correlation is significant at the .01 level (2-tailed).

Table 9: VIF for each Predictor

Variable	VIF
STEMID	3.622
Interest	15.912
Recognition	2.319
Performance	5.537
Future Self	6.261
STEMID_Score	66.836

Tests of the assumption of collinearity indicated that there was evidence of multicollinearity between STEMID_Score and Interest. The Pearson correlations between STEMID_Score and Interest is well above the "rule of thumb" value of .80 ($r(334) = .909, p = .01$) and the VIF of STEMID_Score is well above 4 ($VIF = 66.836$). Because of this multicollinearity, the fact that STEMID_Score was an included information about all predictors already in the model, and the fact that Interest was our predictor variable of concern, we decided to take STEMID_Score out of the model, after which we saw no evidence of multicollinearity.

Table 10: Final Model Results. PISEC is the outcome variable (1=yes; 0=no)

Predictor	b	SE(b)	Wald	OR	$CI_{95\%}$ for OR		p-value
					Lower	Upper	
STEMID	.290	.338	.739	1.337	.690	2.591	.390
Recognition	.320	.392	.668	1.378	.639	2.970	.414
Performance	.932	.594	2.464	2.540	.793	8.135	.116
Future Self	.137	.372	.135	1.147	.553	2.378	.713
Interest	1.683	.736	5.235	5.382	1.273	22.751	.022
Constant	-10.691	2.110	25.670	.000			

Table 10 above shows the results of our final model. In order to assess the fit quality of our model, we used two goodness of fit (GoF) tests: likelihood ratio test [36] and the Hosmer & Lemeshow goodness of fit test [37]. For the likelihood ratio test the statistical hypotheses are:

H_0: $b_i = 0$, i = 1, 2, 3, 4, and 5

H_A: At least one b_i not equal to 0

According to our analyses of the model results: $D_0 = 260.866$; $D_M = 218.993$, so $G^2 = D_0 – D_M = 40.873$. Based on this result, we conclude that the model containing the six predictors fits the data better than the null model ($G^2 = 40.873, df = 5, p < .001$).

For the Hosmer and Lemeshow test the statistical hypotheses are:

H_0: My model adequately fits the data.

H_A: No, it does not.

From the results our model: $X^2 = 2.947$, df = 8, p = .938. This result indicates that our model containing the five predictors adequately fits the data ($X^2 = 3.093, df = 8, p = .928$).

In order to assess the amount of variance explained in our model, we used the likelihood ratio R^2_L for the model ($R^2_L = 1 – (D_M/D_0)$). R^2_L is useful here because it treats D_M like $SS_{Residual}$ and D_0 like SS_{Total} from OLS regression [38]. In this way R^2_L is similar in nature to R^2_{OLS}, and because of the way it is constructed, it can be interpreted as the proportion deviance accounted for. Using the D_0 and D_M values from above gives $R^2_L = 1 – (218.993/260.866)$, so $R^2_L = .161$.

Our model fits the data adequately based on the two GoF tests we used. If we were to assess the model by the R^2_L metric, we could reach a similar conclusion. This is because our results indicate that a model with only 5 predictors for an outcome that is as nuanced as the decision to participate in an informal STEM education program behavior explains roughly 16% of the deviance in our model. We believe that this provides evidence that interest plays a significant role in determining whether or not a person will choose to participate in ISE.

The results shown in Table 10 above suggest that of the affective indicators examined in this study, Interest is most critical in predicting whether or not a youth will choose to participate in the PISEC program. Given the strong collinearity between interest and our total STEMID_score (see above), this result makes sense, even after removing STEMID_score from our model.

We can see from Table 10 that $OR_{Interest} = 5.382$, which indicates that a youth who initially has high STEM interest has 5 times the odds of choosing to participate in the PISEC program compared to a youth who initially has lower STEM interest. This implies that there is a STEM interest bias among the students recruited into PISEC. Understanding this finding lays the groundwork for PISEC and other informal STEM program stakeholders to better understand ways to attract those who are not initially interested in STEM into their programs so they can serve as a conduits for change in STEM within local communities.

Another striking feature that can be seen in Table 10 is that the confidence interval for interest is quite large ($CI_{Interest} = 21.478$). For this study, a wide confidence interval indicates that there is a greater amount of uncertainty about the impact of STEM interest on the decision to participate PISEC. This is reasonable because the model only contains 5 variables on an outcome that could be impacted by many other factors. While this study's results are consistent with prior literature, and are promising, a more detailed model is needed to more fully understand the factors associated with the decision to participate in ISE, and to better understand the impact interest has on this decision. This study was an initial step in understanding this important topic, and the authors are currently working to refine their study in order to improve the scholarly community's understanding of the types of students currently attracted to informal STEM education programs.

One last feature that we found interesting about the results of our study is that STEMID, Recognition, and Future Self are roughly independent of the decision to participate in PISEC ($OR_{STEMID} = 1.337$, $OR_{Recognition} = 1.378$, $OR_{Future\_Self} = 1.147$). This may suggest that some of the factors that constitute STEM identity are not necessarily present when a youth begins participating in an informal STEM education; and that those traits form as a result of their participation. It could be that STEM interest, while being a component of STEM identity, is also a primary initiator of STEM identity formation, with the other factors being developed as one becomes more deeply engaged in STEM. Further study is necessary to explore this idea further, however this work provides initial evidence that ISE performs a key role in shaping the ability of youth to see themselves as scientists and engineers. This further establishes the importance of attracting youth with lower STEM interest into ISE programming, particularly youth from low income and racially marginalized backgrounds.

Discussion and Conclusions

Regarding RQ1, our study found that primary-aged youth who chose to participate in PISEC had significantly higher initial STEM interest than similarly-aged youth in our control group. This suggests that PISEC is attracting youth who are have initial STEM interest. This finding is important for parents, youth and PISEC stakeholders.

The program's model provides STEM interested youth with three key elements necessary for STEM identity formation [10], [39]:

1. opportunities to engage in hands-on STEM experiences via PISEC activities;
2. access to STEM mentors; and
3. early STEM exposure via PISEC participation.

It stands to reason that these programmatic elements keep youth who have high interest in STEM from falling through the cracks. Given that many PISEC's partner schools primarily consist of underrepresented racially diverse youth, this ISE program can serve as a platform for fostering equity in STEM. Further study is needed to understand how the program can best position itself to broaden participation in STEM.

The first tenet of phase I of the four-phase model of interest posits that in the Triggered Situation Interest Phase, interest can be initiated by environmental factors such as personal relevance [14]. This tenet is quite useful in understanding our findings from RQ1. Youth initially become interested in participating in ISE programs, such as PISEC, because they see it as having personal relevance. This personal relevance starts with initial STEM interest, which is coupled with other factors such as: interest in doing STEM activities; having something to do after school; parental influence; opportunities to hang out with friends and peers. Our RQ1 finding suggests that initial STEM interest can motivate ISE participation. Our framework posits that once interest in program participation is triggered, having the programmatic elements allows ISE programs to play the important role of giving youth a platform that allows them to continually re-engage in STEM content, which as they advance through the other three phases of interest, can help them form a positive STEM identity.

In our investigation of RQ2, after controlling for participants' self-reported sense of STEM identity, performance, recognition, and future-self, our study found a statistically significant relationship between primary-aged youth's initial STEM interest and their decision to participate in the PISEC program. Primary-aged youth with high levels initial STEM interest were around 5 times more likely to choose to participate in the program than similarly-aged youth with low initial STEM interest.

It is important to note our hypothesized relationship between initial STEM interest and STEM identity formation. There is consensus among researchers that STEM identity is formed over time [40], [41]. This result has two important implications. Viewing our finding through the lens of our framework suggests that while initial interest may motivate youth participation in ISE, ISE programs must work to keep them engaged in order to allow their progression through all four phases of interest.

We know from prior research that ISE participation positively impacts STEM interest and STEM identity formation [9], [42]. This means that ISE programs are currently engaging in work that transforms their participants' initial STEM interest into fully formed interest and ultimately, positive STEM identity formation. What is less understood is the process that takes youth from initial interest to having positive STEM identity formation. Further research is needed to understand the longitudinal impacts that ISE has on youth. The findings presented here are part of a larger study designed to examine the longitudinal impacts of ISE university-community partnerships on youth's STEM identity formation.

As is true of any research, this study has several limitations. First, this study's sample was derived from a multisite investigation of one program and its partner schools. While the findings represent an addition of new fundamental knowledge to pre-college engineering education and ISE scholarship, further study is needed across multiple contexts. IRB restrictions prevented us from considering demographics in our study. This means that important factors such as race and socioeconomic status were absent from our analysis. We intend to include such demographics in our future work.

The strong relationship between initial STEM interest and PISEC participation suggests that ISE programs may be primarily attracting youth who are already interested in STEM. While further study is needed, this may suggest a need by ISE stakeholders to look at ways they can attract youth with low initial STEM interest as this could enhance their ability to increase and broaden participation in STEM.

Acknowledgements

This work is supported by funding from the JILA PFC: NSF PHY-2317149. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Nation Science Foundation. We would like to thank the PISEC school and community partners, youth participants, and university volunteers who helped make this research possible.

References

[1] National Science Board, “Science and Engineering Indicators 2018,” National Science Foundation, Alexandria, VA, NSB-2018-1, 2018.

[2] *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. Washington, D.C.: National Academies Press, 2007. doi: 10.17226/11463.

[3] S. Olson and D. Riordan, “Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics,” Report to the Office of the President, Feb. 2012. Accessed: Apr. 16, 2018. [Online]. Available: https://files.eric.ed.gov/fulltext/ED541511.pdf

[4] K. P. Dabney *et al.*, “Out-of-School Time Science Activities and Their Association with Career Interest in STEM,” *Int. J. Sci. Educ. Part B*, vol. 2, no. 1, pp. 63–79, Mar. 2012, doi: 10.1080/21548455.2011.629455.

[5] P. Bell, B. Lewenstein, A. W. Shouse, and M. A. Feder, *Learning Science in Informal Environments: People, Places, and Pursuits*, vol. 140. Washington, DC: National Academies Press, 2009. Accessed: Oct. 13, 2020. [Online]. Available: http://www.tandfonline.com/doi/full/10.1179/msi.2009.4.1.113

[6] C. D. Denson, C. Hailey, C. A. Stallworth, and D. L. Householder, “Benefits of Informal Learning Environments: A Focused Examination of STEM-based Program Environments,” vol. 16, no. 1, p. 5, 2015.

[7] A. Bicer and Y. Lee, “Effect of STEM PBL embedded informal learning on student interest in STEM majors and careers,” *J. Math. Educ.*, vol. 12, no. 1, pp. 57–73, 2023, doi: https://doi.org/10.26711/007577152790038.

[8] M. W. Kier, M. R. Blanchard, J. W. Osborne, and J. L. Albert, “The Development of the STEM Career Interest Survey (STEM-CIS),” *Res. Sci. Educ.*, vol. 44, no. 3, pp. 461–481, Jun. 2014, doi: 10.1007/s11165-013-9389-3.

[9] R. Dou, Z. Hazari, K. Dabney, G. Sonnert, and P. Sadler, “Early informal STEM experiences and STEM identity: The importance of talking science,” *Sci. Educ.*, vol. 103, no. 3, pp. 623–637, May 2019, doi: 10.1002/sce.21499.

[10] J. Henderson, V. Snodgrass Rangel, J. Holly, R. Greer, and M. Manuel, “Enhancing Engineering Identity Among Boys of Color,” *J. Pre-Coll. Eng. Educ. Res. J-PEER*, vol. 11, no. 2, Sep. 2021, doi: 10.7771/2157-9288.1311.

[11] N. S. King and R. M. Pringle, “Black girls speak STEM: Counterstories of informal and formal learning experiences,” *J. Res. Sci. Teach.*, vol. 56, no. 5, pp. 539–569, May 2019, doi: 10.1002/tea.21513.

[12] H. B. Carlone and A. Johnson, “Understanding the science experiences of successful women of color: Science identity as an analytic lens,” *J. Res. Sci. Teach.*, vol. 44, no. 8, pp. 1187–1218, Oct. 2007, doi: 10.1002/tea.20237.

[13] A. V. Maltese, C. S. Melki, and H. L. Wiebke, “The nature of experiences responsible for the generation and maintenance of interest in STEM,” *Sci. Educ.*, vol. 98, no. 6, pp. 937–962, 2014, doi: 10.1002/sce.21132.

[14] S. Hidi and K. A. Renninger, “The four-phase model of interest development,” *Educ. Psychol.*, vol. 41, no. 2, pp. 111–127, Jun. 2006, doi: 10.1207/s15326985ep4102_4.

[15] R. Hughes, B. Nzekwe, and K. Molyneaux, “The Single Sex Debate for Girls in Science: A Comparison Between Two Informal Science Programs on Middle School Students’ STEM Identity Formation,” *Res. Sci. Educ.*, vol. 43, Oct. 2013, doi: 10.1007/s11165-012-9345-7.

[16] Z. Hazari, G. Sonnert, P. M. Sadler, and M.-C. Shanahan, “Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study,” *J. Res. Sci. Teach.*, p. n/a-n/a, 2010, doi: 10.1002/tea.20363.

[17] M. J. Graham, J. Frederick, A. Byars-Winston, A.-B. Hunter, and J. Handelsman, “Increasing Persistence of College Students in STEM,” *Science*, vol. 341, no. 6153, pp. 1455–1456, 2013.

[18] A. C. Barton and E. Tan, “*We Be Burnin’!* Agency, Identity, and Science Learning,” *J. Learn. Sci.*, vol. 19, no. 2, pp. 187–229, Apr. 2010, doi: 10.1080/10508400903530044.

[19] A. Calabrese Barton, H. Kang, E. Tan, T. B. O’Neill, J. Bautista-Guerra, and C. Brecklin, “Crafting a Future in Science: Tracing Middle School Girls’ Identity Work Over Time and Space,” *Am. Educ. Res. J.*, vol. 50, no. 1, pp. 37–75, Feb. 2013, doi: 10.3102/0002831212458142.

[20] H. B. Carlone *et al.*, “‘Unthinkable’ Selves: Identity boundary work in a summer field ecology enrichment program for diverse youth,” *Int. J. Sci. Educ.*, vol. 37, no. 10, pp. 1524–1546, Jul. 2015, doi: 10.1080/09500693.2015.1033776.

[21] R. Elmesky, “"I Am Science and the World Is Mine^^i Embodied Practices as Resources for Empowerment,” *Sch. Sci. Math.*, vol. 105, no. 7, pp. 335–342, 2005.

[22] K. Riedinger and J. McGinnis, “An investigation of the role of learning conversations in youth’s authoring of science identities during an informal science camp,” *Int. J. Sci. Educ. Part B*, vol. 7, pp. 1–27, Jan. 2017, doi: 10.1080/21548455.2016.1173741.

[23] K. P. Dabney, R. H. Tai, and M. R. Scott, “Informal Science: Family Education, Experiences, and Initial Interest in Science,” *Int. J. Sci. Educ. Part B*, vol. 6, no. 3, pp. 263–282, Jul. 2016, doi: 10.1080/21548455.2015.1058990.

[24] X. Kong, K. P. Dabney, and R. H. Tai, “The Association Between Science Summer Camps and Career Interest in Science and Engineering,” *Int. J. Sci. Educ. Part B*, vol. 4, no. 1, pp. 54–65, Jan. 2014, doi: 10.1080/21548455.2012.760856.

[25] K. P. Dabney, T. N. Johnson, G. Sonnert, and P. M. Sadler, “STEM career interest in women and informal science,” *J. Women Minor. Sci. Eng.*, vol. 23, no. 3, 2017, doi: 10.1615/JWomenMinorScienEng.2017018018.

[26] B. Habig, P. Gupta, B. Levine, and J. Adams, “An Informal Science Education Program’s Impact on STEM Major and STEM Career Outcomes,” *Res. Sci. Educ.*, vol. 50, no. 3, pp. 1051–1074, Jun. 2020, doi: 10.1007/s11165-018-9722-y.

[27] C. Maiorca *et al.*, “Informal Learning Environments and Impact on Interest in STEM Careers,” *Int. J. Sci. Math. Educ.*, vol. 19, no. 1, pp. 45–64, Jan. 2021, doi: 10.1007/s10763-019-10038-9.

[28] K. Miller, G. Sonnert, and P. Sadler, "The influence of students' participation in STEM competitions on their interest in STEM careers," *Int. J. Sci. Educ. Part B*, vol. 8, no. 2, pp. 95–114, Apr. 2018, doi: 10.1080/21548455.2017.1397298.
[29] J. L. Young, J. R. Young, and D. Y. Ford, "Culturally Relevant STEM Out-of-School Time: A Rationale to Support Gifted Girls of Color," *Roeper Rev.*, vol. 41, no. 1, pp. 8–19, Jan. 2019, doi: 10.1080/02783193.2018.1553215.
[30] H. B. Carlone, C. M. Scott, and C. Lowder, "Becoming (less) scientific: A longitudinal study of students' identity work from elementary to middle school science: BECOMING (LESS) SCIENTIFIC," *J. Res. Sci. Teach.*, vol. 51, no. 7, pp. 836–869, Sep. 2014, doi: 10.1002/tea.21150.
[31] F. Herrera, S. Hurtado, G. Garcia, and J. Gasiewski, "A Model for Redefining STEM Identity For Talented STEM Graduate Students," Jun. 2005.
[32] A. Johnson, J. Brown, H. Carlone, and A. K. Cuevas, "Authoring identity amidst the treacherous terrain of science: A multiracial feminist examination of the journeys of three women of color in science," *J. Res. Sci. Teach.*, vol. 48, no. 4, pp. 339–366, Apr. 2011, doi: 10.1002/tea.20411.
[33] M. M. Williams and C. E. George-Jackson, "Using and Doing Science: Gender Self-Efficacy, and Science Identity of Undergraduate Students in STEM," *J. Women Minor. Sci. Eng.*, vol. 20, no. 2, pp. 99–126, 2014, doi: 10.1615/JWomenMinorScienEng.2014004477.
[34] Z. Zacharia and A. C. Barton, "Urban middle-school students' attitudes toward a defined science," *Sci. Educ.*, vol. 88, no. 2, pp. 197–222, Mar. 2004, doi: 10.1002/sce.10110.
[35] A. Agresti, *An introduction to categorical data analysis*, 2nd ed. in Wiley series in probability and mathematical statistics. Hoboken, NJ: Wiley-Interscience, 2007.
[36] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*, Third edition. in Wiley series in probability and statistics. Hoboken, New Jersey: Wiley, 2013.
[37] D. W. Hosmer and S. Lemesbow, "Goodness of fit tests for the multiple logistic regression model," *Commun. Stat. - Theory Methods*, vol. 9, no. 10, pp. 1043–1069, Jan. 1980, doi: 10.1080/03610928008827941.
[38] T. Tjur, "Coefficients of Determination in Logistic Regression Models—A New Proposal: The Coefficient of Discrimination," *Am. Stat.*, vol. 63, no. 4, pp. 366–372, 2009.
[39] N. A. Ortiz, T. R. Morton, M. L. Miles, and R. S. Roby, "What About Us? Exploring the Challenges and Sources of Support Influencing Black Students' STEM Identity Development in Postsecondary Education," *J. Negro Educ.*, vol. 88, no. 3, pp. 311–326, 2019.
[40] L. Chu *et al.*, "Argument-Driven Engineering in Middle School Science: An Exploratory Study of Changes in Engineering Identity Over an Academic Year," *J. Pre-Coll. Eng. Educ. Res. J-PEER*, vol. 9, no. 2, Oct. 2019, doi: 10.7771/2157-9288.1249.
[41] K. L. Tonso, "Teams that Work: Campus Culture, Engineer Identity, and Social Interactions," *J. Eng. Educ.*, vol. 95, no. 1, pp. 25–37, 2006, doi: 10.1002/j.2168-9830.2006.tb00875.x.
[42] Lee, Hyonyong, Kwon, Hyuk-Soo, and Park, Kyung-Suk, "Student Motivation and Interests as Proxies for Forming STEM Identities," *J. Korean Assoc. Sci. Educ.*, vol. 32, no. 3, pp. 532–540, Jun. 2012, doi: 10.14697/JKASE.2012.32.3.532.