Strengthening the STEM pipeline: impact of project-based synthetic biology program on high school students' science identity and competency

Pamela J. Mims[1]*, Lindsay E. Lee[1], Natalie Kuldell[2] and Chloe Franklin[2]

[1]Department of Educational Foundations and Special Education, East Tennessee State University, Johnson City, TN, United States, [2]BioBuilder Educational Foundation, Boston, MA, United States

This study investigates the impact of a project-based science education intervention, BioBuilderClub, on high school students' science identity, self-beliefs, and content knowledge in synthetic biology. Addressing the critical "leaky pipeline" issue in biotechnology education, this intervention focused on fostering scientific engagement and competency through hands-on, interdisciplinary projects. Using descriptive and correlational statistics (i.e., paired *t*-tests, residual change regression), we found that the project-based intervention resulted in significant improvements in students' self-perceived scientific engagement, competency, and content knowledge regardless of gender, locale, and first generation status. Across expert raters, we also found an improvement in the understanding of synthetic biology and students reported an increased interest in biotechnology and related fields. These findings underscore the potential of project-based learning to enhance STEM retention by building strong science identities, particularly among underrepresented groups. Future research should explore the long-term impacts of such interventions and their integration into standard curricula to further bolster the biotechnology pipeline.

KEYWORDS

project-based teaching, synthetic biology, STEM education, STEM career awareness, research

Introduction

Synthetic biology's focus on creative bio-design processes and real-world applications makes it a promising field for educational interventions aimed at addressing the "leaky pipeline" issue, where students drop out of the pathway from high school to post-secondary education and careers in science, technology, engineering, and math (STEM) fields. The COVID-19 pandemic has renewed calls to improve student retention in the STEM pipeline by providing access to high-quality, equitable STEM education opportunities both inside and outside of school. The pandemic underscored the importance of scientific advancement and a robust STEM workforce while exposing the education system's struggle to recover lost ground (White House, 2022). Recent educational initiatives, including those by the U.S. Department of Education, seek to address various factors—both cognitive and non-cognitive—that influence students' decisions to remain in STEM pathways at critical stages of their education (Marten, 2022; U.S. Department of Education, 2023).

Many students who initially express an interest in pursuing STEM careers lose that intention during high school. For example, a longitudinal study of over 24,000 high school students revealed a 48% decrease in those intending to pursue STEM careers between 9th and 11th grade (Mangu et al., 2015). Since students' post-secondary STEM choices are closely tied to their high school science experiences and self-perceptions, educational interventions that provide meaningful science experiences and foster positive self-perceptions may enhance retention and equity within the STEM pipeline (Wang, 2013; National Science Board and National Science Foundation, 2021).

Career trajectories in STEM fields and the factors influencing related decisions are complex and dynamic. However, students with positive self-perceptions related to science—those who identify as a "science person"—are more likely to persist in STEM pathways (Metcalf, 2010; Vincent-Ruz and Schunn, 2018; Chen et al., 2021). Science identity, particularly among underrepresented student populations in STEM, is one of the strongest predictors of persistence in studying STEM subjects throughout high school and college, as well as commitment to a STEM career (Aschbacher et al., 2010; Barton et al., 2013; Chemers et al., 2011; Le et al., 2014). Understanding how educational interventions can promote science identity is crucial for educators, policymakers, and STEM professionals to positively impact students' self-perceptions and, ultimately, their persistence in STEM pathways.

In human development, identity refers to an individual's "sense of self," which is shaped by their participation in specific activities, societal roles, group affiliations, interests, and personal characteristics (Burke and Stets, 2009; Renninger, 2009). Science identity, being both social and contextual, involves a student's self-perception of "who they are, what they believe they are capable of, and what they want to do and become with regard to science" (Brickhouse, 2001; Kim et al., 2018). According to the model described by Carlone and Johnson (2007), science identity comprises three dimensions: competence, performance, and recognition. Competence refers to possessing and understanding scientific knowledge and methods; performance involves demonstrating this competence to others; and recognition involves being acknowledged as a "science person" by oneself and by peers and significant others. Educational interventions that foster a sense of community and affiliation, promote positive attitudes, and align school science activities with authentic scientific practices have been shown to enhance science identity among adolescents (Vincent-Ruz and Schunn, 2018). A strong science identity can empower students to view themselves as active and valued contributors to science, rather than passive learners.

One strategy that has shown some promise for increasing students' science identity is Project-based learning (PBL). PBL is an educational strategy that holds significant potential for enhancing science identity and increasing STEM retention, particularly among underrepresented groups. PBL emphasizes collaborative, student-centered learning through projects that are meaningful, relevant, and connected to real-world challenges. Research has shown that PBL fosters deeper engagement, improves problem-solving skills, and cultivates a sense of ownership over learning (Krajcik and Shin, 2014; Mergendoller and Thomas, 2013). Among underrepresented groups, PBL can provide a culturally responsive approach to STEM education by incorporating students' lived experiences and interests into the learning process, thereby enhancing their sense of belonging and self-efficacy in science (Barton et al., 2013; Ladson-Billings, 1995). Additionally, PBL supports the development of science identity by allowing students to see themselves as active participants in the scientific process, gaining recognition for their contributions and developing confidence in their scientific abilities (Chin and Osborne, 2008; Moote et al., 2020).

Building on these general benefits of PBL, synthetic biology and biotechnology education provide fertile ground for applying PBL principles, offering authentic, interdisciplinary challenges that can further enhance science identity and engagement. Synthetic biology offers unique opportunities for educational interventions aimed at fostering science identity, while also benefiting from the increased diversity and strength of a more robust STEM pipeline. Educational interventions in synthetic biology often involve interdisciplinary projects that challenge students to tackle real-world problems from diverse contexts and to evaluate or create biological solutions using engineering design thinking (iGEM Foundation, 2024; NISE Network, 2024). These interventions, which integrate life sciences with the engineering design-build-test-learn cycle, highlight creativity, innovation, real-world relevance, and authenticity—factors that are known to enhance science identity and increase student persistence in STEM fields (Shanahan and Nieswandt, 2009; Vincent-Ruz and Schunn, 2018). Engaging high school students in authentic, original research has been linked to increased research identity, motivation, and a higher likelihood of pursuing and sustaining a career in STEM, compared to students who do not engage in original research until college (Deemer et al., 2022). Synthetic biology research interventions at the secondary education level may be especially crucial in addressing the "leaky pipeline," as they infuse science education with elements that promote science identity at a critical juncture for STEM student retention (Graham et al., 2013). Walker (2021) conducted a survey of 66 middle school students, followed by interviews, to explore what this age group knows and thinks about synthetic biology and modern biotechnology applications. The study found that while students had minimal knowledge in these areas, their attitudes were similar to those of high school students, suggesting that these attitudes can be shaped during their remaining years of schooling. Walker also highlighted the need for further research to support student learning throughout their educational journey, emphasizing the importance of developing a strong pipeline to meet the growing demand in these fields. Additionally, Walker's analysis of prior research indicates that factors such as age, science learning experiences, and the geopolitical landscape significantly influence students' knowledge of and interest in biotechnology.

In light of Walker (2021) findings, there is a need for additional research on effective instructional formats for teaching synthetic biology content. Project-based science interventions introduced at earlier educational stages may be beneficial, but research is required to assess their overall effectiveness. One such project-based science intervention is the BioBuilderClub, an extracurricular program offered by the BioBuilder Educational Foundation, which emphasizes authenticity in synthetic biology education. The BioBuilderClub engages high school students worldwide in experiential learning, allowing them to tackle real-world challenges with local or global significance. Students work virtually with a practicing bioengineer mentor as they design biological solutions using the design-build-test-learn engineering cycle. While some teams advance to the build and test phases, the BioBuilderClub season culminates in a celebration of student achievements. Here, all teams have the opportunity to present their bio-design processes and research findings through various formats, including oral reports, poster presentations, written abstracts, and publication in peer-reviewed journals.

Purpose

To date, there is a limited number of studies investigating the impact of project-based science education interventions like BioBuilderClub on student outcomes. Consequently, this study aimed to assess the effectiveness of such an intervention on high school students' science self-beliefs and content knowledge over time (i.e., 2023–2024). Such research could guide teachers' use of evidence based approaches for teaching STEM education in high schools. The study was guided by the following research questions:

1. How do high school students perceive their scientific engagement, competency, and content knowledge after participating in a project-based after-school science program?

2. To what extent does students' content knowledge of synthetic biology change after their participation in a project-based after-school science program?

Materials and methods

Prior to the study, institutional review board approval was obtained. The following study used multiple methods to assess the effectiveness of the project-based science education intervention. Specifically, we used a quasi-experimental one-group pre-posttest design to assess students' engagement, competency, and content knowledge before and after the project-based science program. We explored the factor structure and internal consistency of our developed survey on scientific

TABLE 1 Demographic background frequency counts.

	Pre 23–24 (n = 246)	Post 23–24 (n = 137)	Pre/Post 23–24 (n = 124)
Gender			
Female	142 (58%)	79 (58.1%)	73 (58.9%)
Male	92 (37.6%)	52 (38.2%)	46 (37.1%)
Non-binary	1(0.4%)	–	1(0.8%)
My answer is not listed here	3(1.2%)	1 (0.7%)	1(0.8%)
Prefer not to answer	7(2.9%)	4 (2.9%)	3(2.4%)
Race			
Asian (South Asian, Asian American, Pacific Islander)	104 (42.4%)	59 (43.3%)	53 (42.7%)
Black or African American	10 (4.08%)	–	5 (4%)
Hispanic or Latino	12 (4.8%)	5 (3.6%)	8 (6.5%)
White (origins in Europe, the Middle East, or North Africa)	75 (30.6%)	43 (31.6%)	39 (31.5%)
Two or more races	–	–	9 (7.3%)
Other	26 (10.6%)	21 (15.4%)	–
Prefer not to say	18 (7.3%)	8 (5.8%)	10 (8.1%)
Locale (NCES)			
City (Large)	44 (17.9%)	13 (9.5%)	13 (10.5%)
City (Midsize)	14 (6.7%)	8 (5.9%)	8 (6.5%)
City (Small)	18 (7.3%)	14 (10.3%)	12 (9.7%)
Suburb (Large)	164 (66.9%)	101 (74.2%)	91(73.4)
Rural (Distant)	1 (0.4%)	–	–
Rural (Fringe)	4 (1.6%)	–	–
Grade			
9th	9 (3.7%)	6 (4.4%)	4 (3.2%)
10th	32 (13.1%)	19 (14%)	16 (12.9%)
11th	110 (44.9%)	64 (47.1%)	59 (47.6%)
12th	94 (38.4%)	47 (34.6%)	45 (36.3%)
First generation status			
No	107 (43.7%)	61 (44.9%)	54 (43.5%)
Prefer not to say	30 (12.2%)	12 (8.8%)	12 (9.7%)
Yes	108 (44.1%)	63 (46.3%)	58 (46.8%)

For student's race/ethnicity, the Pre-22-23 Data included 83 students (11 missing) and the Post 22–23 Data included 21 students (3 missing).

self-beliefs, and we assessed the interrater reliability of the developed rubric of understanding synthetic biology.

Since 2015, over 1,500 students from diverse schools in 22 states have participated in BioBuilderClub. In the 2023–2024 school year, there were a total of 285 students who participated in the BioBuilderClub. From that participation in the 2023–2024 school year, we used a convenience sampling method and included students enrolled to participate in the project-based science program. Both consent by parents and assent from high school students were obtained during the registration process prior to participation in the program. Students were recruited by BiobuilderClub (i.e., program coordinator, instructors) to complete the pre-survey and post-survey online. To maintain anonymity, surveys were de-identified by BioBuilderClub prior to being shared with external evaluators. Students were incentivized to complete the post-survey by entering students into a raffle for an Amazon Gift Card if they completed the post-survey. There were 246 students (86.3%) who completed the pre-survey and 137 students (48.1%) completed the post-survey. However, there were only 124 students (43.5%) who had completed both the pre and post-survey (see Table 1 for participation rates and demographic information).

Measures

The pre- and post-survey included questions to assess the self-perceptions of student scientific engagement, competency, and content knowledge before and after participation in the project-based science program (see Table 2). The development of the survey items was influenced by prior measures, specifically Competency Beliefs in Science (Chung et al., 2016), the Measuring Activation & Engagement (Moore et al., 2011), and the Classroom Undergraduate Research Experience (CURE) Survey (Lopatto, 2009). The quantitative questions used a Likert style format options ranging from: "1 = Strongly Disagree" to "5 = Strongly Agree." The pre-and post-survey also included open-ended responses to indicate specific content knowledge of defining synthetic biology and student experiences in the project-based science program (Table 3).

Data analysis

All quantitative analyses were conducted using R (Version 4.3.1; R Core Team, 2023). We used listwise deletion to remove missing pre and post-test surveys, thus we initially examined all complete paired pre-and-posttest surveys ($n = 124$). Using the lavaan package (Rosseel,

TABLE 2 Likert style survey items.

Scientific competency items (Pre/Post)
I can understand scientific information on websites
I think I am very good at: Giving evidence when I tell my opinion about science
I think I am very good at: Coming up with new ways to solve technical problems in science
I think I am very good at: Coming up with my own science investigations.
Critiquing the scientific work of other students
Scientific engagement items (Pre/Post)
I enjoy learning about science
I engage in science on weekends and during breaks from school.
I would like to know more about jobs that use science.
I will miss studying science when I leave school.
Scientific content knowledge items (Pre/Post)
Working on a lab or project where no one knows the outcome
Working on student-designed projects
Becoming responsible for a part of a project
Reading primary scientific literature
Writing a research proposal
Collecting data
Analyzing data
Presenting scientific ideas or results orally (spoken presentation)
Presenting scientific ideas or results in written papers or reports
Presenting scientific posters
Maintaining a laboratory notebook
Experience items (Post)
I would recommend a high school class that has BioBuilder content over one that does not
BioBuilder has supported my interest in pursuing advanced scientific coursework
BioBuilder has supported my interest in pursuing a career in the life sciences
In my college selection process, I will be looking for places with biotechnology, biomanufacturing, or synthetic biology opportunities.
BioBuilder has increased the importance of finding biotechnology, biomanufacturing, or synthetic biology opportunities in my college selection process.
In my future job search, I will be looking for places with biotechnology, biomanufacturing, or synthetic biology opportunities.
BioBuilder has increased the importance of finding biotechnology, biomanufacturing, or synthetic biology opportunities in my future job search.

The item of "Critiquing the scientific work of others" was initially thought of as a scientific content knowledge item but showed more relation to scientific competency (see Table 4).

2012), exploratory factor analysis was conducted using oblique rotation (Promax), which is appropriate for examining correlated factors (Tabachnick and Fidell, 2001). We assessed eigenvalues, scree plots, and factor loadings to evaluate a two and three-factor solution.

For the initial pre-and-posttest quantitative analyses, we assessed item-level data. Using paired samples t-tests, we initially conducted descriptive statistics to investigate mean score differences from each item of the pre-post survey responses with complete cases ($n = 124$). Descriptive statistics were used to summarize frequencies among demographics, measures of central tendency, and variability of demographics on the items used to create the factors of scientific engagement, competency, and content knowledge before/after participation in the BioBuilder program. We also reported mean difference scores, percentage difference, and Cohen's d as a measure of effect size (Cohen, 1988; see Table 1).

TABLE 3 Pre-posttest open-ended questions.

Pre-post questions	Type of question
What is synthetic biology?	Content knowledge
Provide an example of synthetic biology. You may provide more than one example.	Content knowledge

Residual change score differences

Using the composite factors (i.e., Scientific Competency, Engagement, and Content Knowledge), we investigated the relationship between pre and post scores with residual change score regressions. Residual change regressions assesses alterations in the outcome (i.e., post-test) from a prior occasion (i.e., pre-test), with other predictors of interest. Castro-Schilo and Grimm (2018) state "the autoregressive effect *residualizes* the outcome leaving only variability that is unexplained by [the pre-test], which can be construed as the variability due to change (see Castro-Schilo and Grimm, 2018, p. 36). Prior studies have explored residualized change regressions to understand longitudinal outcomes of participation in interventions over time (Allen et al., 1994; Kang, 2022) and others have compared its' usage to traditional ANCOVA and difference scores (Jennings and Cribbie, 2016). To further investigate the relation of covariates (i.e., first-generation status, gender, NCES Locale designations) to post scores on three different outcome variables (i.e., Competency, Engagement, & Content Knowledge Beliefs), we conducted three multiple regressions to assess residual change on only complete cases with demographic variables of interest ($n = 103$). Prior to analyses, we used listwise deletion to use complete cases of demographics of interest (i.e., gender, locale, first generation
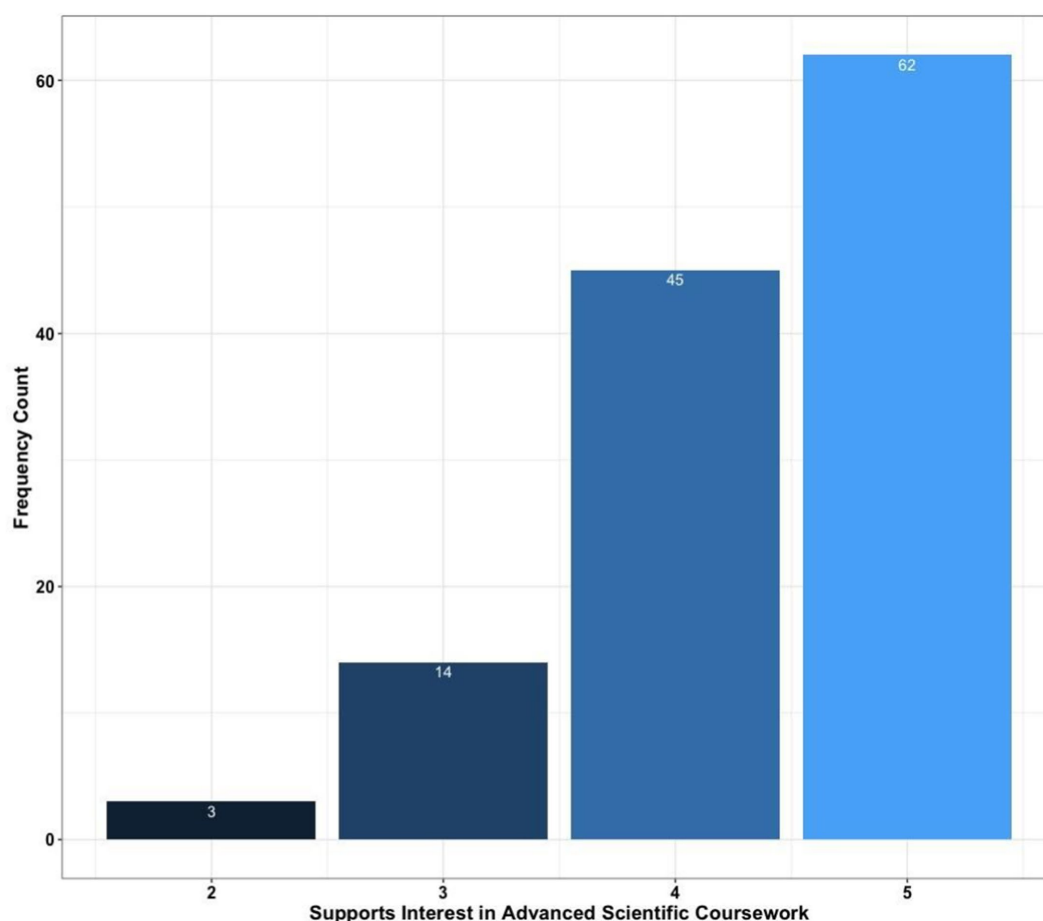

FIGURE 1
Experience: recommendation of BioBuilder.
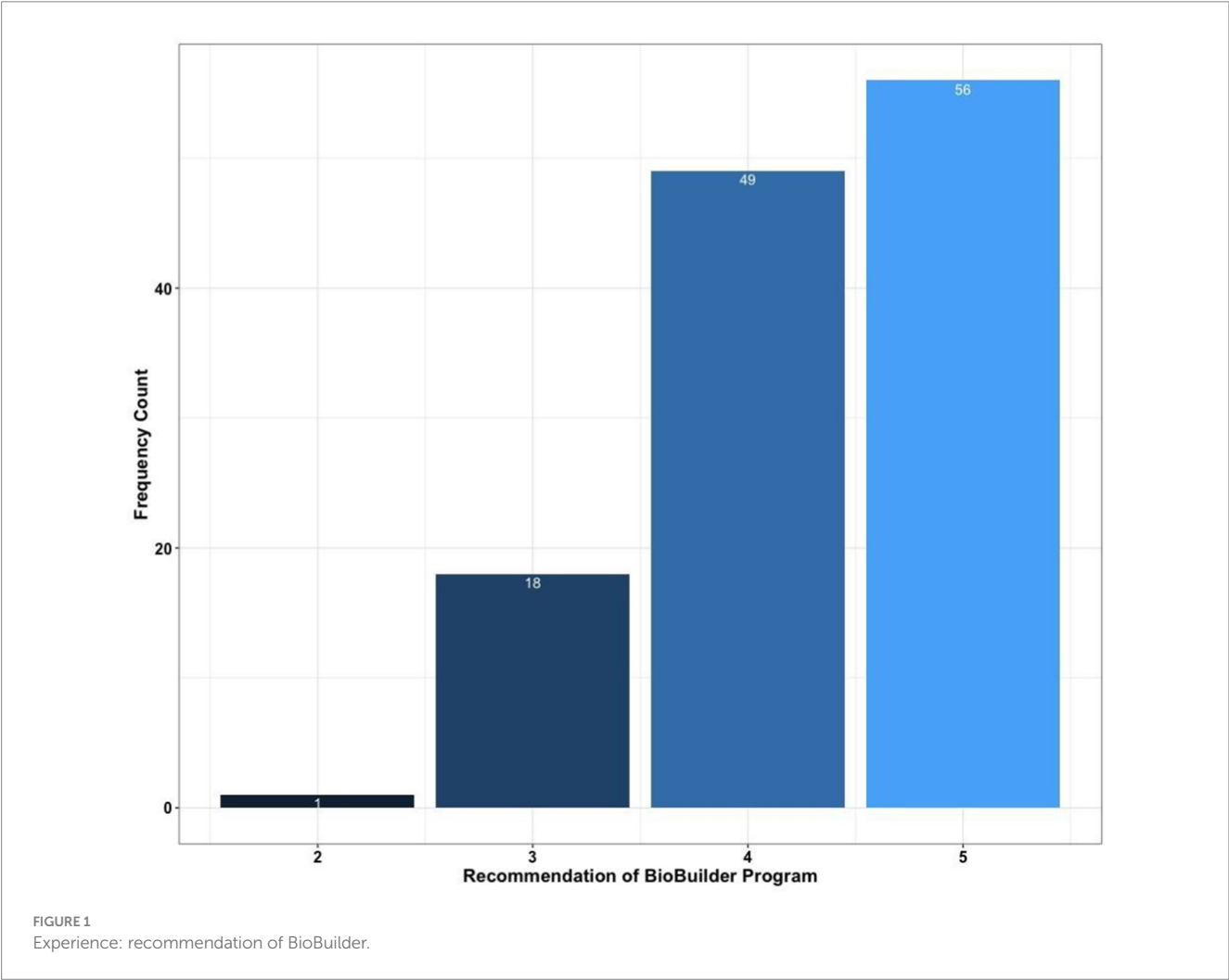
FIGURE 2
Experience: supported my interest in pursuing advanced scientific coursework.

status) and specifically did not include international students from large cities ($n = 7$), students who did not report if they were first-generation ($n = 12$), or students who chose a gender other than male or female designations (i.e., non-binary [$n = 1$], my answer is not listed [$n = 1$], prefer not to answer [$n = 3$]), as there were only a few cases. There was overlap in students omitting responses within these designations. We used a Benjamini Hochberg correction to account for multiple comparisons and provide conservative estimates to avoid possible Type I error (Benjamini and Hochberg, 1995).

Synthetic biology content knowledge: definition and examples

The accuracy of the student responses to content knowledge questions (e.g., what is synthetic biology, provide examples of synthetic biology; see Tables 2, 3) was assessed with a specific rubric. To assess the content knowledge of the question "what is synthetic biology?," we developed a rubric to reflect poor, emerging, or sophisticated responses. Two content expert raters, external to the research team, were asked to rate student submissions. We went over the rubric individually with each rater before they rated submissions. The first content expert rater rated all 124 submissions. The second content expert rater rated a

random sampling of 20% of the 124 submissions. To determine agreement, we calculated interrater nominal agreement and Cohen's Kappa (Cohen, 1960). Using the randomly sampled 20% of pre and post scores, we report means and standard deviations for each rater to show changes from pre to post scores of content knowledge.

Recommendations, college selection, and future job searches

Using a 1-to-5 Likert scale from strongly disagree to strongly agree, we asked students a series of conclusive survey items regarding their experiences and recommendations of the BioBuilder Program, as well as their interest in pursuing biotechnology, biomanufacturing, or synthetic biology in college selection or job searches. We reported frequencies of student responses (see Figures 1–7).

Results

For the 2023–2024 school year, there were 286 students enrolled in the project-based science program. Of that total 286, 246 individuals responded to the pre-survey and 137 individuals responded to the
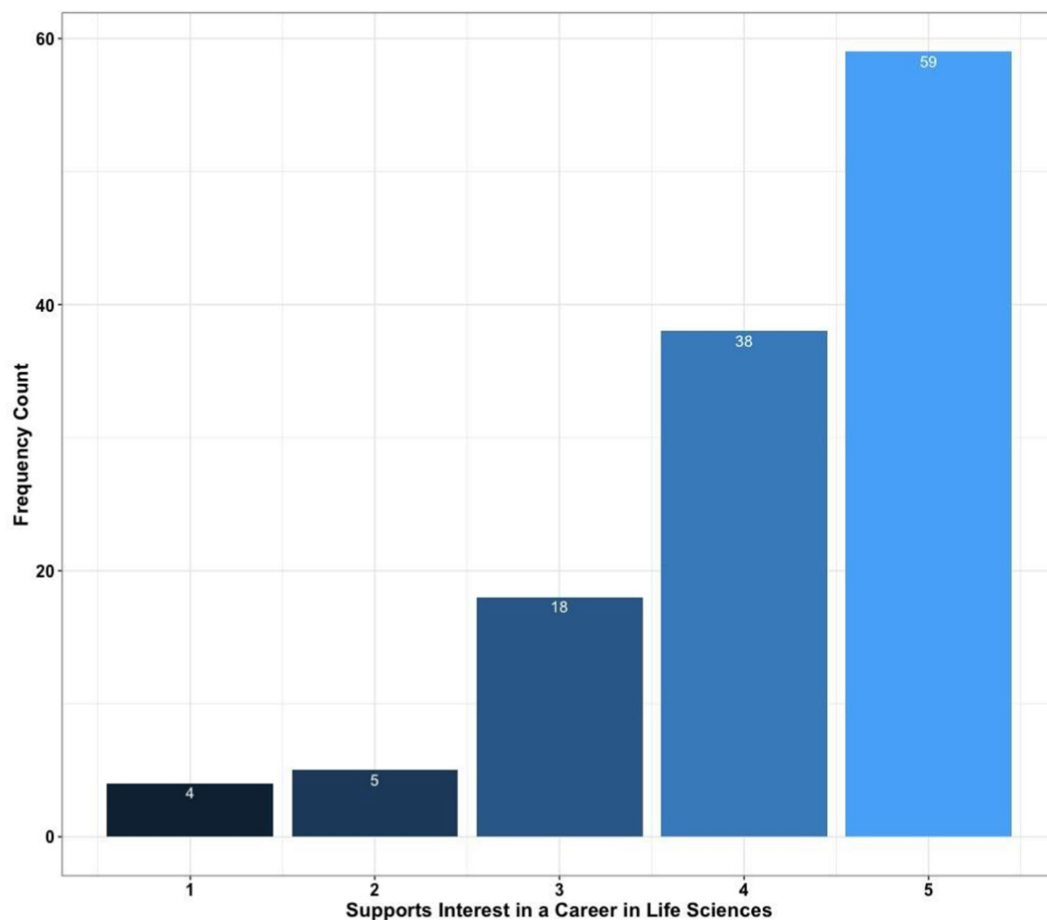
FIGURE 3
Experience: supported my interest in pursuing a career in the life sciences.

post-survey. Of the 137 students, there was a higher response rate for the same individuals who completed both the pre-and post-survey ($n = 124$). For the frequency of demographics across pre and post survey results in the 2023–2024 School Year (see Table 1).

Exploratory factor analysis and reliability

We examined the factor structure of the items in the survey related to Engagement, Competency, and Content Knowledge (Table 4). The three-factor model showed some indication of good fit, X^2 (133, 124) = 247.9, $p < 0.001$, and Root Mean Square Error of Approximation (RMSEA = 0.08), but the root mean square of the residuals (RMSR = 0.06) and the Tucker Lewis Index (TLI = 0.83) were not within goodness of fit thresholds (see Kline, 2015). However, both the factor loadings (see Figure 8) and the Parallel Analyses Scree Plot showed an indication of a three-factor model (see Figure 9). Thus, we continued with a three-factor model. One item asking about Critiquing the Scientific Work of Others showed indication of loading within the factor related to Scientific Competency rather than Scientific Content Knowledge (as originally intended). We named these factors: Scientific Engagement (4 items), Scientific Competency (5 items), and Scientific Content Knowledge (11 items). Standard internal consistency estimates for Scientific Engagement ($\alpha = 0.79$),

Scientific Competency ($\alpha = 0.80$), and Content Knowledge ($\alpha = 0.91$) were found to be in the range of acceptable reliability.

Mean difference scores

Initially, we analyzed the mean score differences of the pre-post survey items and found some of the specific scientific competency, engagement, and content knowledge showed higher differences across pre-post.

Scientific competency

For students' competency beliefs, we found students scored higher on the following items: understanding scientific websites, giving evidence when telling their opinion, coming up with new ways to solve a problem, and critiquing the scientific work of other students. For instance, we found students, after the project-based science program, scored higher on their self-beliefs of being able to understand scientific information on websites ($M = 4.04$, $SD = 0.68$) than before they participated in the program ($M = 3.82$, $SD = 0.86$), $t(123) = 3.31$, $p < 0.01$. We found a mean score difference of 0.23 and had a small effect ($d = 0.30$). Also, after participation in the project-based science program, students scored higher on their self-belief of giving evidence when telling
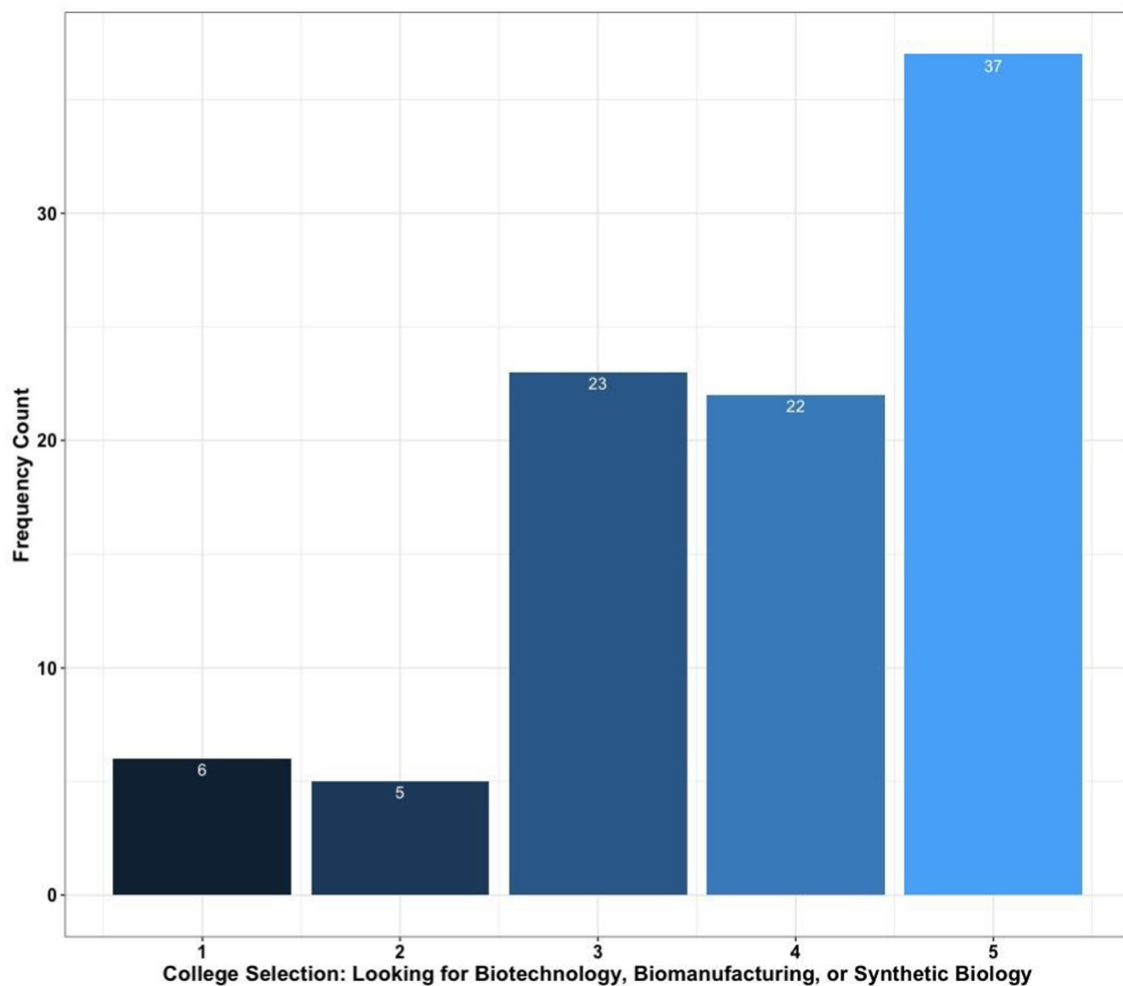
FIGURE 4
Experience: college selection process—looking for places with biotechnology, biomanufacturing, or synthetic biology opportunities.

their opinion about science ($M = 3.9$, $SD = 0.79$), than before they participated in the program ($M = 3.74$, $SD = 0.93$), $t(123) = 2.17$, $p < 0.05$; This showed a mean score difference of 0.16 and showed a minimal effect ($d = 0.18$). After participation students also rated their self-belief of coming up with new ways to solve technical problems ($M = 3.66$, $SD = 0.81$) than before participation in the project-based science program ($M = 3.46$, $SD = 0.90$), $t(123) = 2.49$, $p < 0.05$. This shows a mean difference of 0.20 and a small effect ($d = 0.24$). They also rated their self-belief of coming up with their own science investigations higher after participation in the project-based science program ($M = 3.65$, $SD = 0.80$), $t(123) = 2.82$, $p < 0.01$. This showed a mean difference of 0.27 and had a small effect ($d = 0.29$).

Scientific engagement

For scientific engagement, students scored higher on engaging in science on the weekends and during breaks from school ($M = 3.8$, $SD = 1.16$), $t(123) = 3.05$, $p < 0.01$. There was a mean score difference of 0.30 and this had a small effect ($d = 0.26$). Students did not have significant differences on enjoyment of learning science, wanting to know more about jobs that use science, and missing science when they leave school. However, all those students' self-beliefs were above average before starting the program.

Scientific content knowledge

For student's beliefs of their scientific content knowledge, we found after the project-based science program students scored higher on working on a lab or project with an unknown outcome ($M = 3.42$, $SD = 1.12$) than before the program ($M = 3.03$, $SD = 1.24$), $t(123) = 3.49$, $p < 0.001$. There was a mean difference of 0.39 and had a small effect ($d = 0.33$). We also found that students scored higher on working on student-design projects ($M = 3.82$, $SD = 1.02$) than before participating in the program ($M = 3.51$, $SD = 1.13$), $t(123) = 2.81$, $p < 0.01$. This had a mean difference of 0.32 and showed a small effect ($d = 0.29$). We also found after participation in the project-based science program, students scored higher on becoming responsible for a project ($M = 4.46$, $SD = 0.77$) than before participating in the program ($M = 4.16$, $SD = 1.02$), $t(123) = 3.04$, $p < 0.01$. This had a mean difference of 0.30 and showed a large effect ($d = 0.33$).

Students' beliefs on specific tasks within the program showed a significant increase after participation in the project-based science program. For instance, we found after participation in program
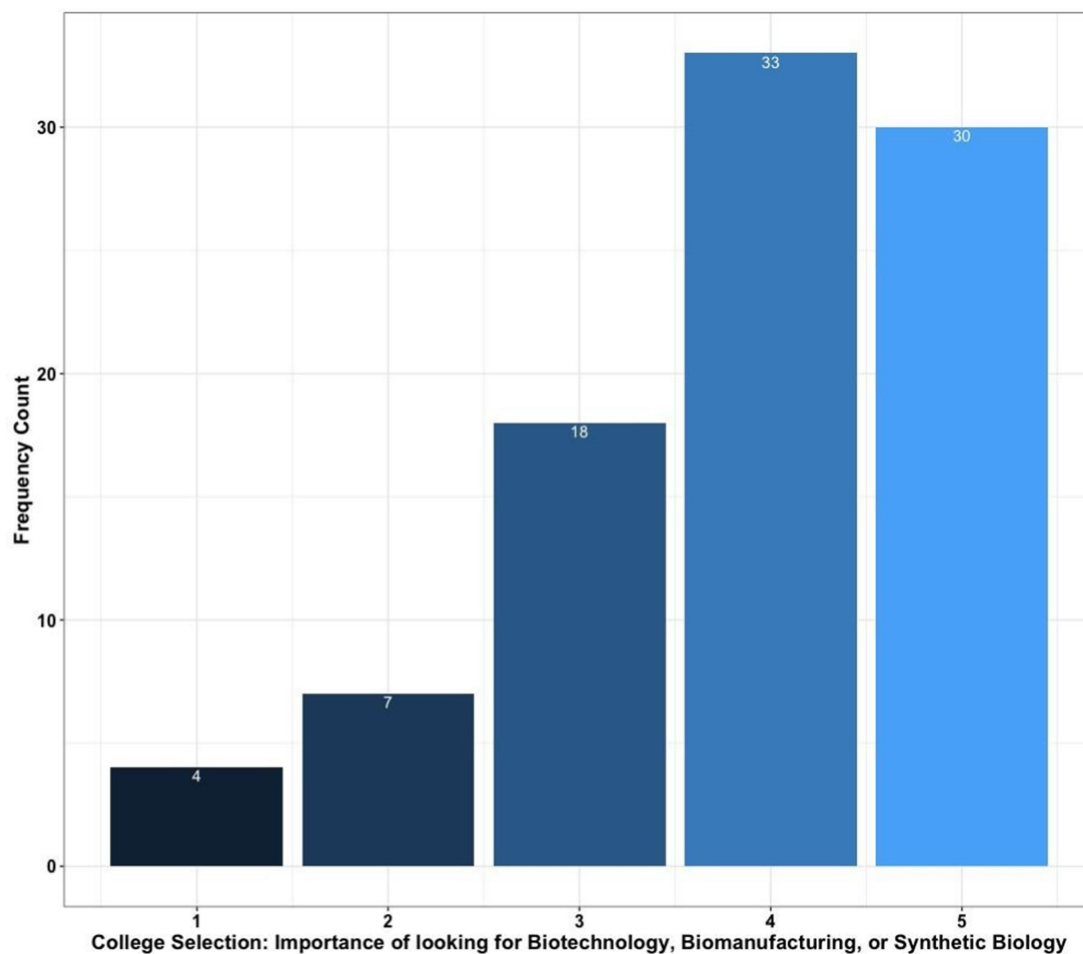
FIGURE 5
Experience: college selection process—importance of finding biotechnology, biomanufacturing, or synthetic biology opportunities.

students scored higher on their self-beliefs of reading primary scientific literature ($M = 3.89$, $SD = 0.94$) than before participating in the program ($M = 3.47$, $SD = 1.14$), $t(123) = 4.17$, $p < 0.001$. This had a larger mean difference of 0.42 and showed a small-medium effect ($d = 0.40$). Students also reported higher on their beliefs of their ability to write a research proposal ($M = 3.49$, $SD = 1.06$) than before participating in the program ($M = 2.99$, $SD = 1.26$), $t(123) = 4.42$, $p < 0.001$. This had the largest mean difference of 0.50 and showed a small-medium effect ($d = 0.43$). Students did not have significant mean differences on how they believed they could collect and analyze data after the duration of the investigation-centered science program.

In terms of externalizing their ideas, we found significant mean differences in their written reports, and creating posters. For instance, after participation in BioBuilder students scored higher on presenting scientific ideas or results in written papers or reports ($M = 3.61$, $SD = 1.03$) than before participating in the program ($M = 3.41$, $SD = 1.20$), $t(123) = 3.06$, $p < 0.01$. This showed a mean difference of 0.32 and showed a small effect ($d = 0.27$). For scientific posters, we found that post-participation in the program, students scored higher on their self-beliefs of presenting a poster ($M = 3.55$, $SD = 1.09$) than before participating in the program ($M = 3.22$, $SD = 1.27$), $t(123) = 2.68$,

$p < 0.01$. This showed a mean difference of 0.33 and a small effect ($d = 0.27$). Students did not have significant mean differences on their self-beliefs of maintaining a laboratory notebook (see Table 5).

Residual change score differences

For the entire sample ($n = 124$), we analyzed descriptive and correlational statistics for pre and post survey scores across composite factors (i.e., Scientific Engagement, Scientific Competency, Scientific Content Knowledge; see Figure 10; Table 6). We found there were moderate to strong relationships between several pre and post scores across factors.

For residual change score differences, we found gender, first generation status, or locale did not significantly predict post-scores across each factor of scientific engagement, scientific competency, and scientific content knowledge (see Table 7). After controlling for each of these covariates, we did find significant residual differences from pre- to post-scores across each factor of scientific engagement, competency, and content knowledge (see Table 6). For instance, we found pre-scores ($b = 0.46$, $p < 0.001$) of students' beliefs of
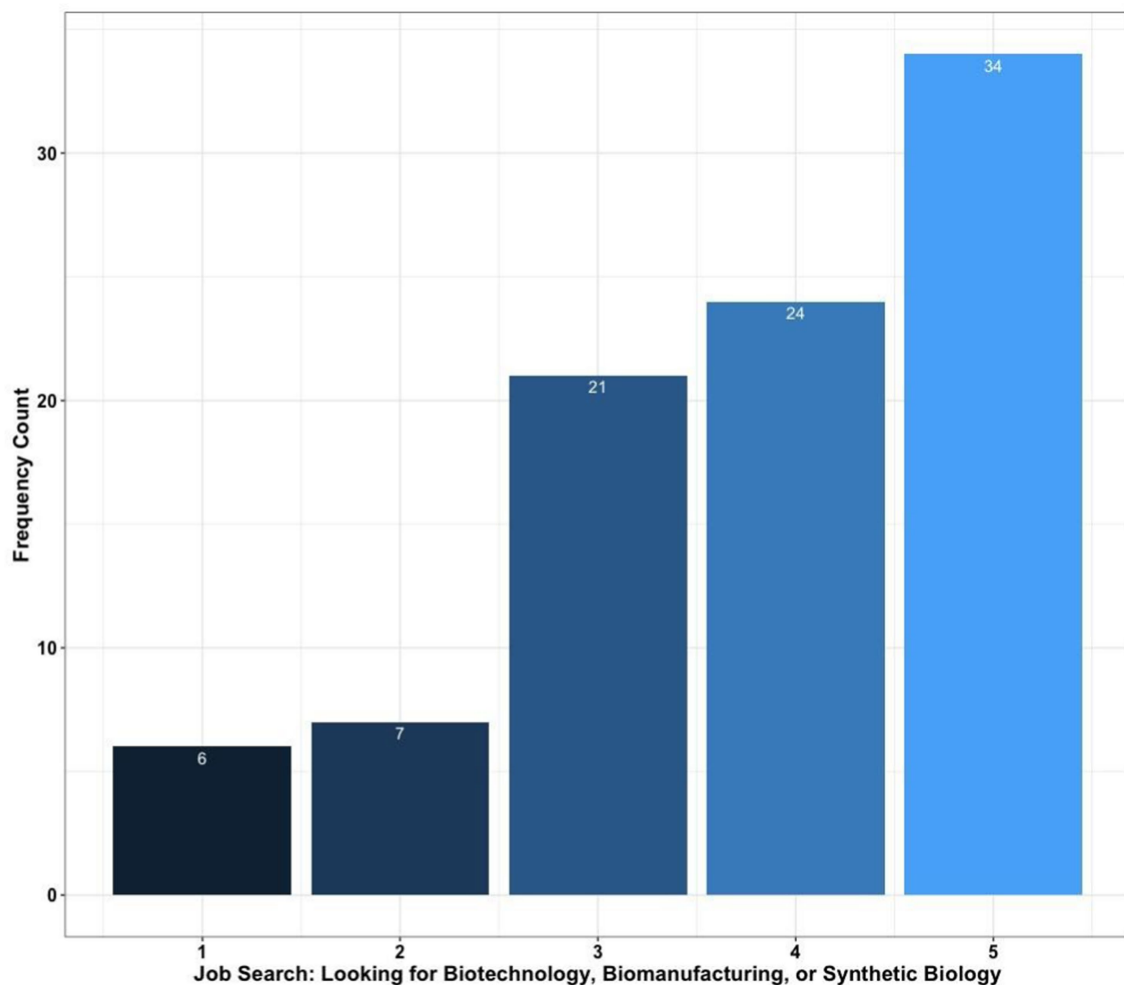
FIGURE 6
Experience: future job search—looking for places with biotechnology, biomanufacturing, or synthetic biology opportunities.

Scientific Content Knowledge did significantly predict post-scores, $R^2 = 0.57$, $F(6, 96) = 13.73$, $p < 0.001$; This means the model accounted for 46% of the variance in students' beliefs of their Scientific Content Knowledge. All three factors of Scientific Engagement, Scientific Competency, Scientific Content Knowledge, showed no significant differences across demographic covariates. Scientific Engagement ($R^2 = 0.46$) and Scientific Content Knowledge ($R^2 = 0.46$) showed the largest effect; however Scientific Competency ($R^2 = 0.36$) also showed a large effect.

Changes in synthetic biology knowledge

Promotor and gene expression

Students were asked if they could confidently identify the symbol for a promoter from a specified figure using a 1-to-5 Likert Scale ($M = 3.7$, $SD = 1.4$). Responses showed a left-skewed distribution (Figure 11). In other words, many students felt confident in identifying the promoter symbol. Students were also asked if the symbols in a figure were in the proper order for a gene expression (Figure 12).

There were 83 students (66.9%) who rated this as TRUE, whereas there were 41 students who said they did not know (26.6%) or chose FALSE (6.5%).

What is synthetic biology?

This first round of rating resulted in poor agreement for both raters across pre-definition responses (37.5%) and post-definition responses (58%), the same was true for pre-example responses (40%) and post-example responses (48%). The researchers and raters met for a consensual discussion of the developed rubric. During the discussion, we discovered that one rater was grading too rigorously, and another was being too lenient. Raters were given the same random sampling of 20% responses with blank scores and asked to separately re-rate submissions after re-interpretation of the rubric (see Tables 8, 9). Across both raters, this second round of rating resulted in slight to moderate agreement across pre-definition responses (68%) and post-definition responses (64%), as well as pre-example responses (72%) and post-example responses (64%).

Although there was improved agreement in the second round, we decided to have a consensual discussion to discuss specific
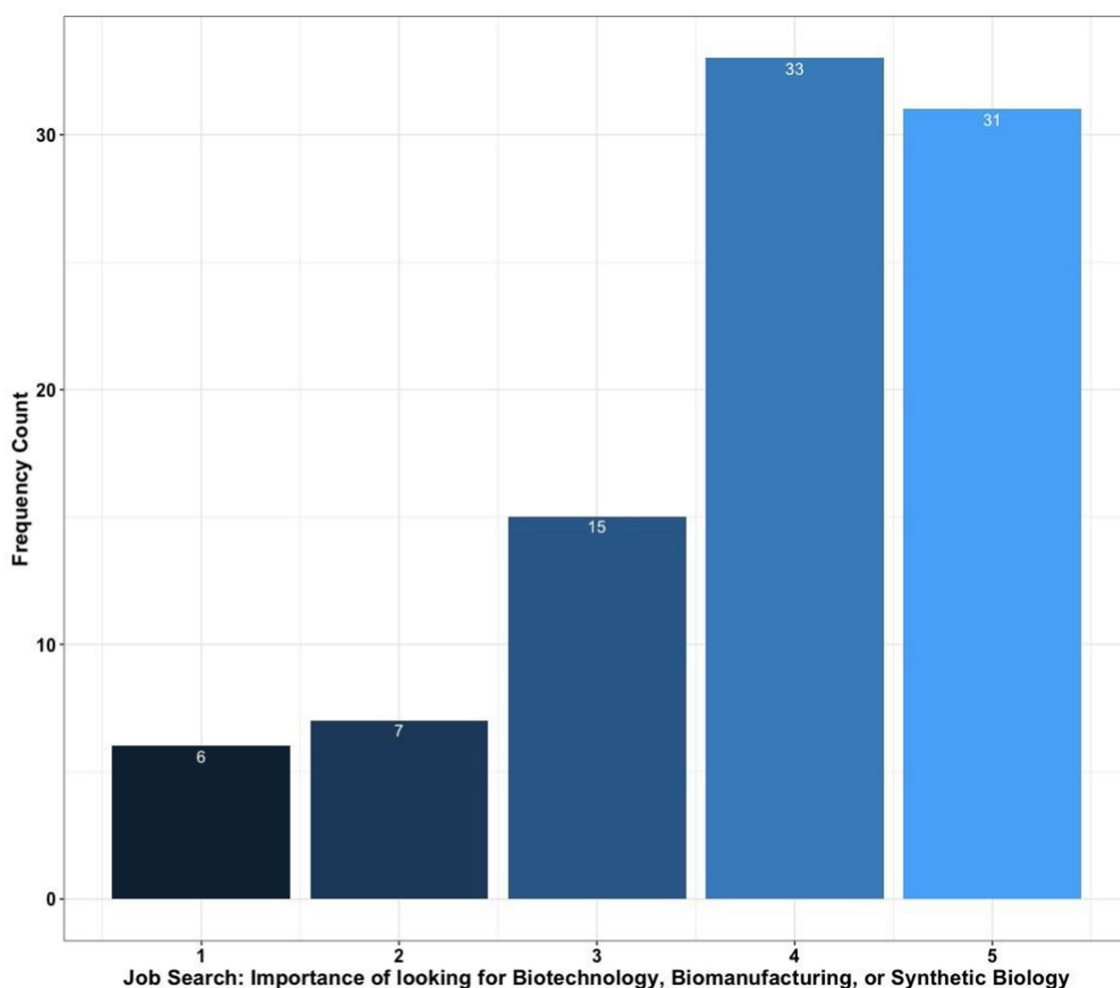
FIGURE 7
Experience: future job search - importance of finding biotechnology, biomanufacturing, or synthetic biology opportunities.

disagreements across raters. We found that Rater 1 was rating post-scores more harshly than Rater 2 because they thought students should have more advanced, nuanced answers after the program. Rater 2 referred to specific language used within the rubric to justify their scores. After this discussion, there were still disagreements, however we found substantially better agreement across pre-definition responses (100%) and post-definition responses (92%), pre-example responses (96%), and post-example responses (88%) (for Cohen's Kappa, see Table 10). Overall, students showed an emerging to sophisticated understanding of synthetic biology from pre-to-post responses, as well as pre-to-post examples (see Figures 13, 14).

Recommendations, college selection, and future job searches

The majority of students "would recommend a high school class that has the project-based science program over one that does not," with a mean score of 4.3 ($SD = 0.7$; see Figure 1). We also asked students if the project-based science program supported their interest in advanced scientific coursework ($M = 4.3$, $SD = 0.8$; Figure 2) or pursuing a career in the Life Sciences ($M = 4.2$, $SD = 1$; Figure 3). For college selection, we asked if students were considering biotechnology, biomanufacturing, or synthetic biology ($M = 3.8$, $SD = 1.1$; Figure 4), as well as the importance of these fields of study for college selection ($M = 3.8$, $SD = 1.2$; Figure 5). When students were looking for future jobs, we asked if they were considering biotechnology, biomanufacturing, or synthetic biology ($M = 3.8$, $SD = 1.2$; Figure 6), as well as the importance of finding a job within these fields ($M = 3.8$, $SD = 1.2$; Figure 7). For all four of these items, students rated them similarly with left skewed distributions indicating more students choosing they would be likely seeking biotechnology, biomanufacturing, or synthetic biology in college selection and future job searches.

Discussion

The findings of this study highlight significant positive outcomes for students participating in the project-based science

TABLE 4 Factor loadings, eigenvalues, and variance (n = 124).

Items	Standardized Factor Loadings (Pattern Matrix)			Communality	Uniqueness	Complexity
	Scientific content knowledge	Scientific competency	Scientific engagement			
I can understand scientific information on websites		**0.72**		0.49	0.51	1
Giving evidence when I tell my opinion about science		**0.63**	0.13	0.47	0.53	1.1
Coming up with new ways to solve technical problems in science		**0.70**		0.50	0.50	1
Coming up with my own science investigations.		**0.56**	0.14	0.37	0.63	1.1
I enjoy learning about science			**0.82**	0.67	0.33	1
I engage in science on weekends and during breaks from school.		0.26	**0.57**	0.49	0.51	1.4
I would like to know more about jobs that use science.			**0.77**	0.59	0.41	1
I will miss studying science when I leave school.		0.17	**0.51**	0.35	0.65	1.2
Working on a lab or project where no one knows the outcome	**0.55**	0.10		0.39	0.61	1.1
Working on student-designed projects	**0.81**	−0.27	0.16	0.48	0.52	1.3
Becoming responsible for a part of a project	**0.79**	−0.21		0.47	0.53	1.2
Reading primary scientific literature	**0.59**		0.13	0.37	0.63	1.1
Writing a research proposal	**0.72**		−0.19	0.54	0.46	1.1
Collecting data	**0.65**			0.41	0.59	1
Analyzing data	**0.66**			0.48	0.52	1
Presenting scientific ideas or results orally (spoken presentation)	**0.66**			0.46	0.54	1
Presenting scientific ideas or results in written papers or reports	**0.56**	0.32	−0.15	0.60	0.40	1.8
Presenting scientific posters	**0.59**	0.27		0.60	0.40	1.5
Critiquing the scientific work of other students	0.36	**0.52**	−0.32	0.60	0.40	2.5
Maintaining a laboratory notebook	**0.34**	0.17		0.22	0.78	1.5
Properties						
SS loadings	4.76	2.60	2.18			
Proportion variance	0.24	0.13	0.11			
Cumulative variance	0.24	0.37	0.48			
Proportion explained	0.50	0.27	0.23			
Cumulative proportion	0.50	0.77	1.00			

Bolded values indicate which items we determined were best explained under each factor (see columns for Scientific Content Knowledge, Scientific Competency, & Scientific Engagement). We first considered moderate correlations higher than a general rule of thumb of 0.32 (Tabachnick and Fidell, 2001). If an item loaded on to two factors, we determined that the item loaded on to the factor with the higher standardized factor loading (i.e., greater than 0.32). See Figure 8 for final model.

education intervention, filling a critical gap in the existing literature regarding the effectiveness of such interventions in enhancing students' science identity, self-beliefs, and content knowledge. Previous studies have emphasized the importance of fostering science identity as a key factor in retaining students in STEM pathways (Aschbacher et al., 2010; Barton et al., 2013; Chemers et al., 2011). This study contributes to that body of knowledge by demonstrating that participation in a project-based
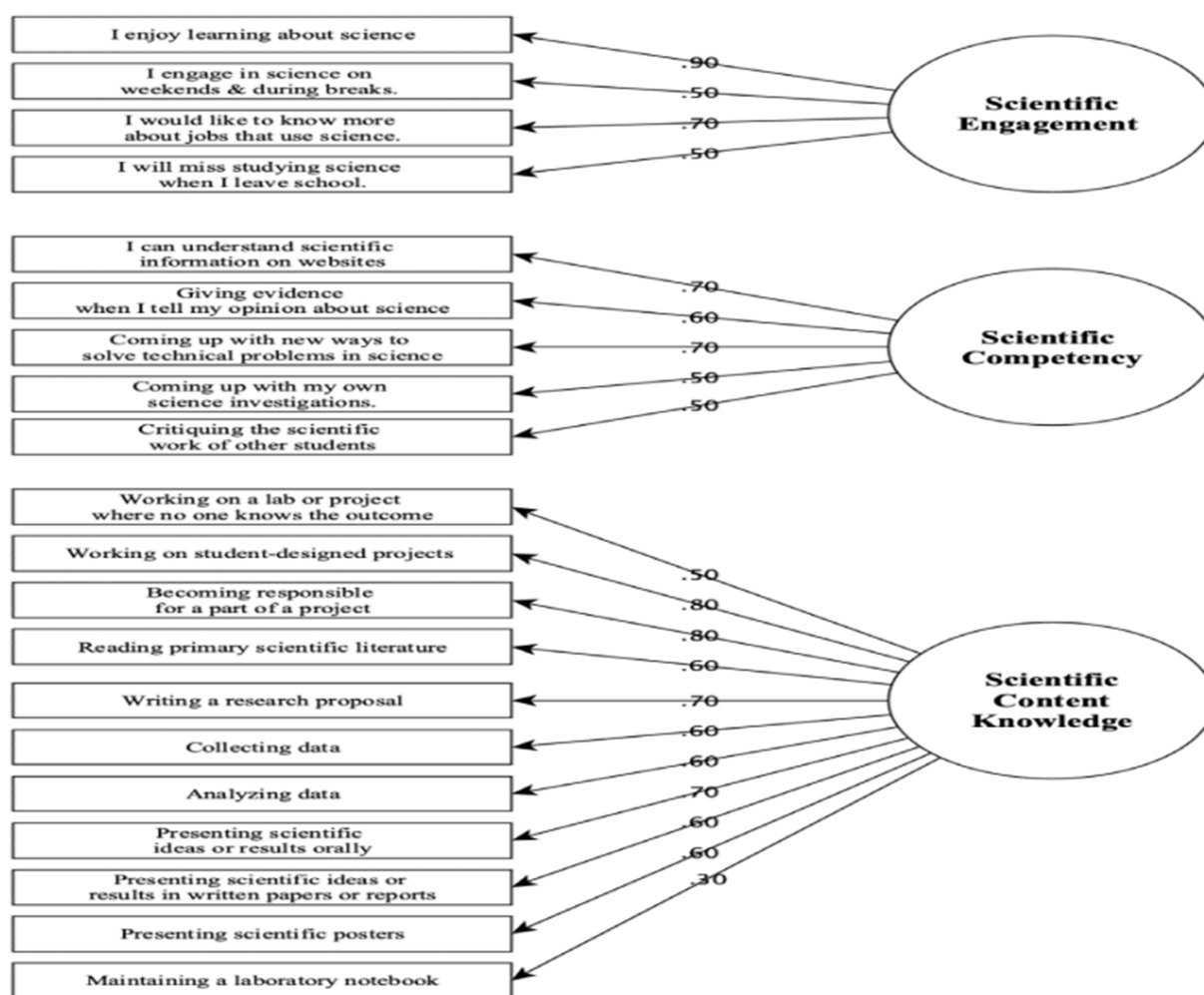
FIGURE 8
Three factor model: scientific engagement, competency, and content knowledge. Standardized factor loadings were rounded to the nearest whole number.
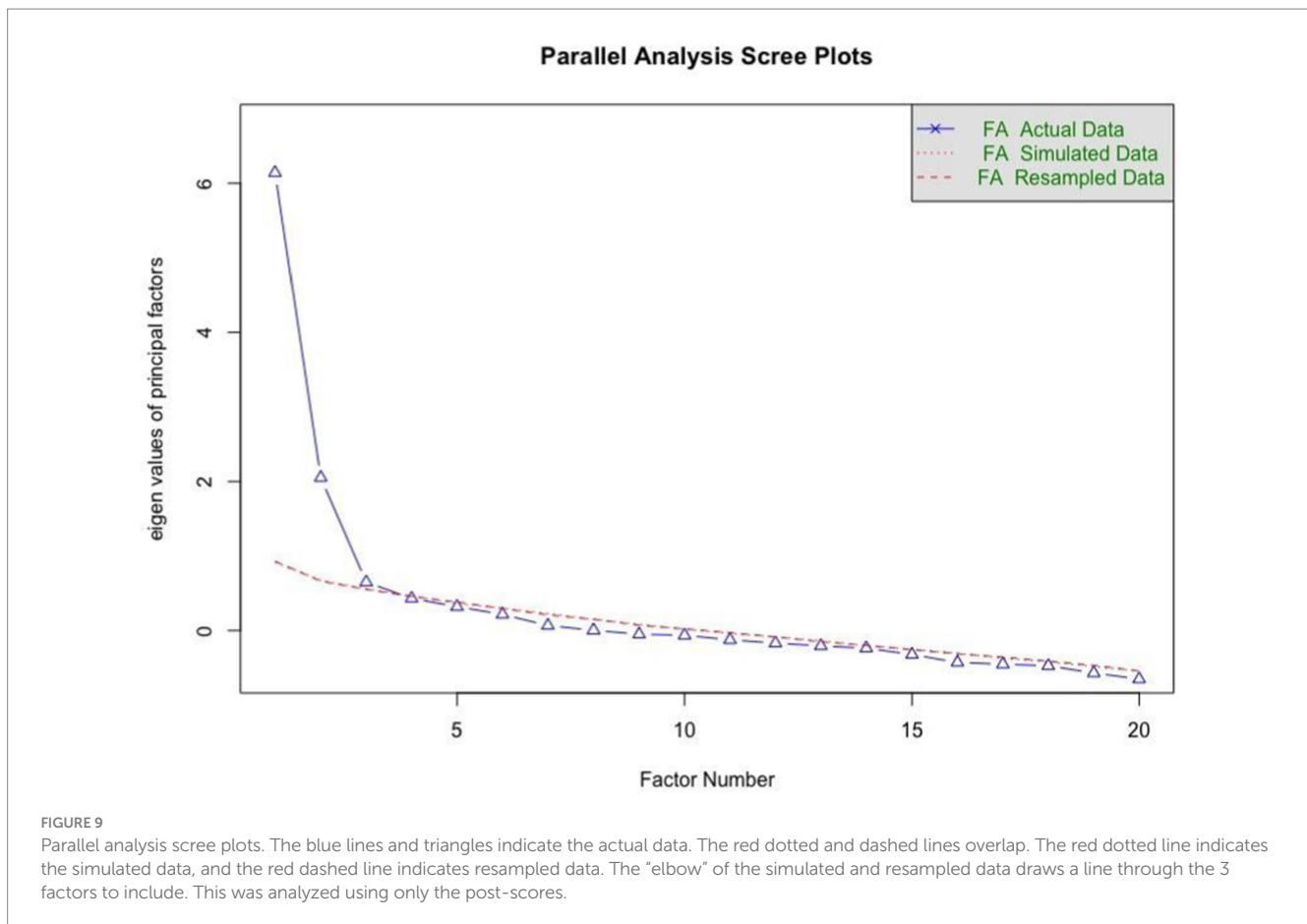
intervention like BioBuilderClub not only increases students' scientific engagement, competency, and content knowledge but also enhances their interest in biotechnology and related fields, aligning with the findings of prior research on the importance of authentic, hands-on learning experiences (Deemer et al., 2022; Shanahan and Nieswandt, 2009).

Our examination of high school students' self-beliefs regarding their scientific engagement, competency, and content knowledge revealed that students who participated in the intervention during the 2023–2024 school year demonstrated significant increases across these areas. This outcome is consistent with the theoretical framework suggesting that students who view themselves as competent and engaged in science are more likely to persist in STEM pathways (Carlone and Johnson, 2007; Vincent-Ruz and Schunn, 2018). The evidence of a three-factor structure in our exploratory factor analysis (Table 4; Figure 8) supports the construct validity of these measures, with internal consistency values ($\alpha = 0.79$–0.91) further confirming the reliability of our findings.

Increased scientific engagement was particularly noteworthy, as students reported greater involvement in scientific activities

beyond the classroom, such as engaging with science during weekends and school breaks. This finding suggests that the project-based science education intervention successfully fosters a deeper interest in science, consistent with literature highlighting the importance of sustained engagement for the development of science identity (Metcalf, 2010; Vincent-Ruz and Schunn, 2018). Additionally, students demonstrated improved scientific competency, including their ability to understand scientific information, generate problem-solving methods, and conduct independent investigations—skills that are critical for future scientific endeavors and align with the competencies described by Carlone and Johnson (2007).

The program also effectively enhanced students' content knowledge in synthetic biology, a field with growing importance in the modern bioeconomy (Graham et al., 2013). Students showed significant gains in understanding key concepts and techniques, such as reading primary scientific literature, writing research proposals, and presenting scientific ideas. This improvement was corroborated by both quantitative measures and qualitative feedback from students, who expressed confidence in their newfound knowledge and skills.

FIGURE 9
Parallel analysis scree plots. The blue lines and triangles indicate the actual data. The red dotted and dashed lines overlap. The red dotted line indicates the simulated data, and the red dashed line indicates resampled data. The "elbow" of the simulated and resampled data draws a line through the 3 factors to include. This was analyzed using only the post-scores.

These outcomes address the gaps identified by Walker (2021) and others who have called for more research on how to effectively teach synthetic biology and related fields at earlier educational stages.

Students overwhelmingly recommended the project-based science education intervention, noting that it provided valuable educational experiences. They particularly appreciated the opportunity to collaborate with peers and mentors, learn in a supportive environment, and engage in creative scientific projects. The program's emphasis on hands-on, practical learning experiences, which has been highlighted as a key factor in successful STEM education (Shanahan and Nieswandt, 2009; Vincent-Ruz and Schunn, 2018), was a critical element of its success.

In addition to the benefits for students, the program aimed to raise awareness of biomanufacturing and modern life sciences among community stakeholders. Engaging parents, teachers, college admissions officers, and relevant businesses is crucial for creating a supportive ecosystem for students pursuing careers in biotechnology, as noted in previous studies (Marten, 2022; U.S. Department of Education, 2023). The program's efforts to communicate and celebrate student achievements publicly helped build a positive perception of biotechnology as a viable and rewarding career path.

However, while students demonstrated increases in engagement, competency, and content knowledge, the residual change regressions for each item ($n = 103$) did not reveal significant differences by gender, locale, or first-generation status. This finding suggests that the intervention was broadly effective across different demographic groups, but it also indicates a need for future research to explore potential differences in outcomes among more diverse student populations, particularly those systematically excluded from STEM fields (Aschbacher et al., 2010; Barton et al., 2013).

For knowledge of synthetic biology, our raters found that students who participated in the intervention demonstrated increased understanding of synthetic biology (see Figures 13, 14). These findings are consistent with the literature on the importance of providing authentic, interdisciplinary science experiences that connect students with real-world problems and solutions (iGEM Foundation, 2024; NISE Network, 2024).

Limitations and future directions

While the findings of this study are promising, caution should be exercised in interpreting the results. The quasi-experimental one-group pre-posttest design, without a control group or random assignment, limits our ability to infer causality. Likewise, using residualized regression without random assignment could have biased estimates (Farmus et al., 2019; Jennings and Cribbie, 2016; Kisbu-Sakarya et al., 2013). Future research should employ more rigorous methodologies, such as randomized controlled trials, and other quantitative analyses to better assess the impact of project-based science education interventions on student outcomes. Likewise, qualitative methods should be explored to understand the specific

TABLE 5 Pre-post intervention paired samples mean score.

Items	Pre 23–24 Mean (SD)	Post 23–24 Mean (SD)	Mean difference	% Change	df	T-test	Cohen's d	CI % 95
Scientific competency items								
I can understand scientific information on websites	3.82 (0.86)	4.04 (0.68)	0.23	6.14%	123	3.31**	0.30	[0.09, 0.37]
Giving evidence when I tell my opinion about science	3.74 (0.93)	3.9 (0.79)	0.16	4.09%	123	2.17*	0.18	[0.01, 0.29]
Coming up with new ways to solve technical problems in science	3.46 (0.90)	3.66 (0.81)	0.20	6.07%	123	2.49*	0.24	[0.04, 0.38]
Coming up with my own science investigations.	3.38 (1.01)	3.65 (0.80)	0.27	7.86%	123	2.82**	0.29	[0.08, 0.45]
Critiquing the scientific work of other students	2.89 (1.23)	3.25 (1.10)	0.36	12.57%	123	2.86**	0.31	[0.11, 0.61]
Scientific engagement items								
I enjoy learning about science	4.63 (0.62)	4.7 (0.59)	0.06	1.22%	123	1.12	0.09	[−0.04, 0.16]
I engage in science on weekends and during breaks from school.	3.5 (1.15)	3.8 (1.16)	0.30	8.52%	123	3.05**	0.26	[0.10, 0.49]
I would like to know more about jobs that use science.	4.22 (0.95)	4.15 (1.06)	−0.08	−1.91%	123	−0.95	0.08	[−0.25, 0.08]
I will miss studying science when I leave school.	3.91 (0.98)	4.00 (0.97)	0.09	2.27%	123	1.03	0.09	[−0.08, 0.26]
Scientific content knowledge items								
Working on a lab or project where no one knows the outcome	3.03 (1.24)	3.42 (1.12)	0.39	13.06%	123	3.49***	0.33	[0.17, 0.62]
Working on student-designed projects	3.51 (1.13)	3.82 (1.02)	0.32	8.94%	123	2.81**	0.29	[0.09, 0.54]
Becoming responsible for a part of a project	4.16 (1.02)	4.46 (0.77)	0.30	6.96%	123	3.04**	0.33	[0.10, 0.48]
Reading primary scientific literature	3.47 (1.14)	3.89 (0.94)	0.42	12.06%	123	4.17***	0.40	[0.22, 0.62]
Writing a research proposal	2.99 (1.26)	3.49 (1.06)	0.50	16.75%	123	4.42***	0.43	[0.28, 0.72]
Collecting data	4.01 (0.95)	4.06 (0.99)	0.05	1.21%	123	0.53	0.05	[−0.13, 0.23]
Analyzing data	3.93 (0.94)	4.07 (0.80)	0.14	3.70%	123	1.68	0.17	[−0.03, 0.32]
Presenting scientific ideas or results orally (spoken presentation)	3.41 (1.20)	3.61 (1.03)	0.20	6%	123	1.93	0.18	[−0.004, 0.41]
Presenting scientific ideas or results in written papers or reports	3.39 (1.22)	3.71 (1.11)	0.32	9.55%	123	3.06**	0.27	[0.11, 0.53]
Presenting scientific posters	3.22 (1.27)	3.55 (1.09)	0.33	10.00%	123	2.68**	0.27	[0.08, 0.56]
Maintaining a laboratory notebook	3.49 (1.24)	3.37 (1.23)	−0.12	−3%	123	−1.26	0.10	[−0.31, 0.07]

There was a total of 124 students who gave both pre/post data in 23–24. We used paired samples t-tests to assess differences in pre/post scores on each item. CI, Confidence Interval. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
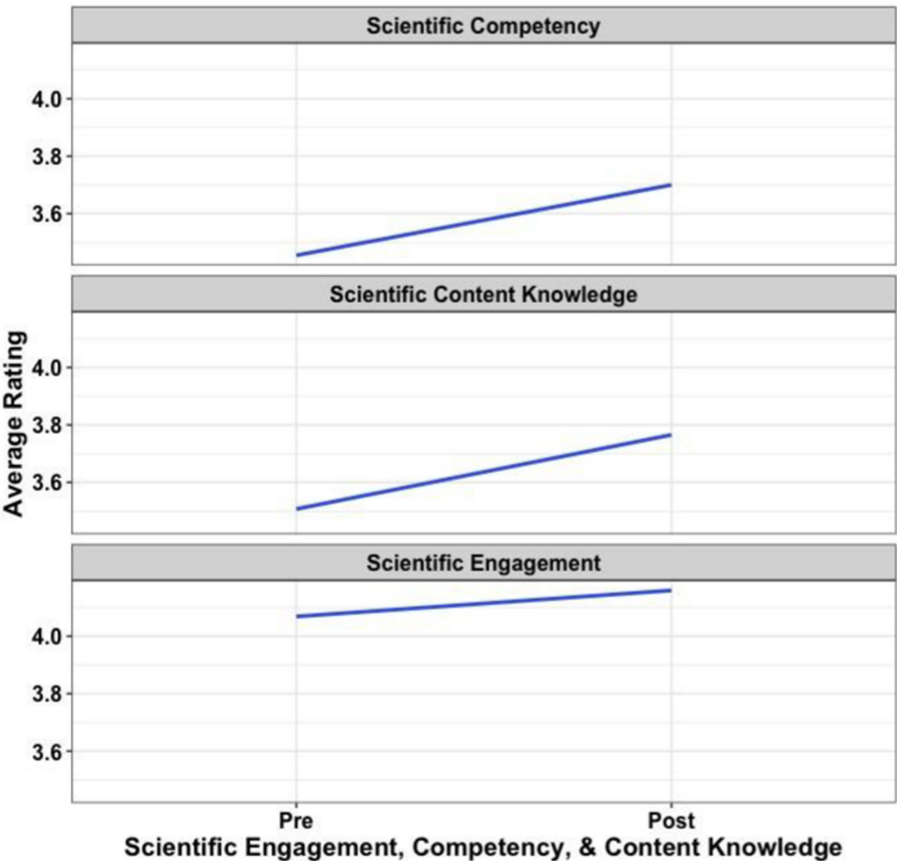
FIGURE 10
Pre-post mean scientific competency, engagement, and content knowledge beliefs ($n = 124$).

TABLE 6 Means, standard deviations, and correlations with confidence intervals.

Variable	M	SD	1	2	3	4	5
1. Scientific engagement (Pre)	4.07	0.73					
2. Scientific competency (Pre)	3.45	0.79	0.56**				
			[0.42, 0.67]				
3. Scientific content knowledge (Pre)	3.51	0.84	0.49**	0.76**			
			[0.34, 0.61]	[0.68, 0.83]			
4. Scientific engagement (Post)	4.16	0.75	0.68**	0.40**	0.37**		
			[0.57, 0.77]	[0.24, 0.54]	[0.20, 0.51]		
5. Scientific competency (Post)	3.70	0.62	0.32**	0.58**	0.49**	0.36**	
			[0.15, 0.47]	[0.45, 0.69]	[0.34, 0.61]	[0.19, 0.50]	
6. Scientific content knowledge (Post)	3.77	0.70	0.14	0.45**	0.60**	0.17	0.60**
			[−0.04, 0.31]	[0.30, 0.58]	[0.48, 0.70]	[−0.00, 0.34]	[0.47, 0.70]

M and SD are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). *$p < 0.05$, **$p < 0.01$.

TABLE 7 Pre-post residual change score differences: competency, engagement, and content knowledge beliefs (n = 103).

Outcome variables	Intercept			Pre			Female			First generation			City: Large			City: Midsize			City: Small			R^2
	B	SE	CI 95%	B	SE	CI 95%	B	SE	CI 95%	B	SE	CI 95%	B	SE	CI 95%	B	SE	CI 95%	B	SE	CI 95%	
Scientific competency	**2.11****	0.24	[1.64, 2.59]	**0.46****	0.07	[0.33, 0.59]	0.05	0.10	[−0.15, 0.25]	−0.03	0.10	[−0.23, 0.17]	−0.12	0.23	[−0.58, 0.34]	0.36	0.20	[−0.04, 0.75]	0.03	0.16	[−0.29, 0.35]	**0.36****
Scientific engagement	**1.53****	0.33	[0.87, 2.19]	**0.64****	0.08	[0.478, 0.80]	0.09	0.11	[−0.13, 0.31]	−0.04	0.11	[−0.26, 0.18]	0.11	0.26	[−0.40, 0.62]	0.40	0.22	[−0.03, 0.83]	0.05	0.17	[−0.30, 0.39]	**0.46****
Scientific content knowledge	**1.80****	0.24	[1.32, 2.28]	**0.57****	0.07	[0.44, 0.71]	0.04	0.11	[−0.18, 0.25]	−0.01	0.11	[−0.01, 0.11]	−0.3	0.24	[−0.91, 0.05]	0.12	0.21	[−0.29, 0.53]	−0.22	0.16	[−0.54, 0.11]	**0.46** ***

There were a total of 124 students who gave both pre/post data in 23–24. For this analysis, we only had 103 students with full demographic information. Residual Change scores show the average population level difference from after participation in the BioBuilder program. The intercept only model should closely resemble the Mean Differences found on Table 5. Residual SE is the standard deviation of the residuals. Smaller residual SE provides indication of greater prediction. Benjamini Hochberg correction to account for multiple comparisons and provide conservative estimates. CI, confidence interval; SE, standard error; Residual SE, residual standard error. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Bolded values indicate statistically significance ($p < 0.05$ or lower).

nuanced lived experiences of students in BioBuilderClub. Additionally, the survey used in this study was a newly developed measure, and only exploratory factor analysis was conducted. Future research should include confirmatory factor analyses to validate the three-factor structure identified and compare it with alternative models using a larger, more diverse sample.

Moreover, while the data were deidentified and sent to independent researchers, the survey development and data collection were conducted by the project-based science education intervention team, which could have influenced the way students responded. Independent investigators should be involved in survey development and data collection in future studies to mitigate potential biases. Additionally, the sample was predominantly White and Asian, with insufficient representation of non-binary students, international students from large cities, and students from systematically excluded racial/ethnic groups (e.g., Black, Hispanic, Native American). Future research should include a more diverse and representative sample to better understand how project-based science education interventions relate to different self-beliefs in scientific engagement, competency, and content knowledge.

The success of the project-based science education intervention has important implications for the future of STEM education. It demonstrates the effectiveness of hands-on, project-based learning in engaging students and enhancing their scientific skills and knowledge. As the demand for skilled workers in the bioeconomy continues to grow, programs like this are essential for preparing the next generation of scientists and engineers. The results of this study support the continued expansion and funding of project-based science education interventions and similar initiatives. Future research should focus on long-term outcomes for participants, including their career trajectories and contributions to the bioeconomy. Additionally, exploring ways to integrate such programs into the standard curriculum could further enhance their impact and reach. By investing in innovative education and training programs, the United States can strengthen its position in the global bioeconomy and ensure a skilled workforce ready to tackle future scientific challenges. Overall, project-based science education interventions represent a promising approach to modernizing biology education and workforce training, with significant potential benefits for students and the broader scientific community.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

PM: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration,
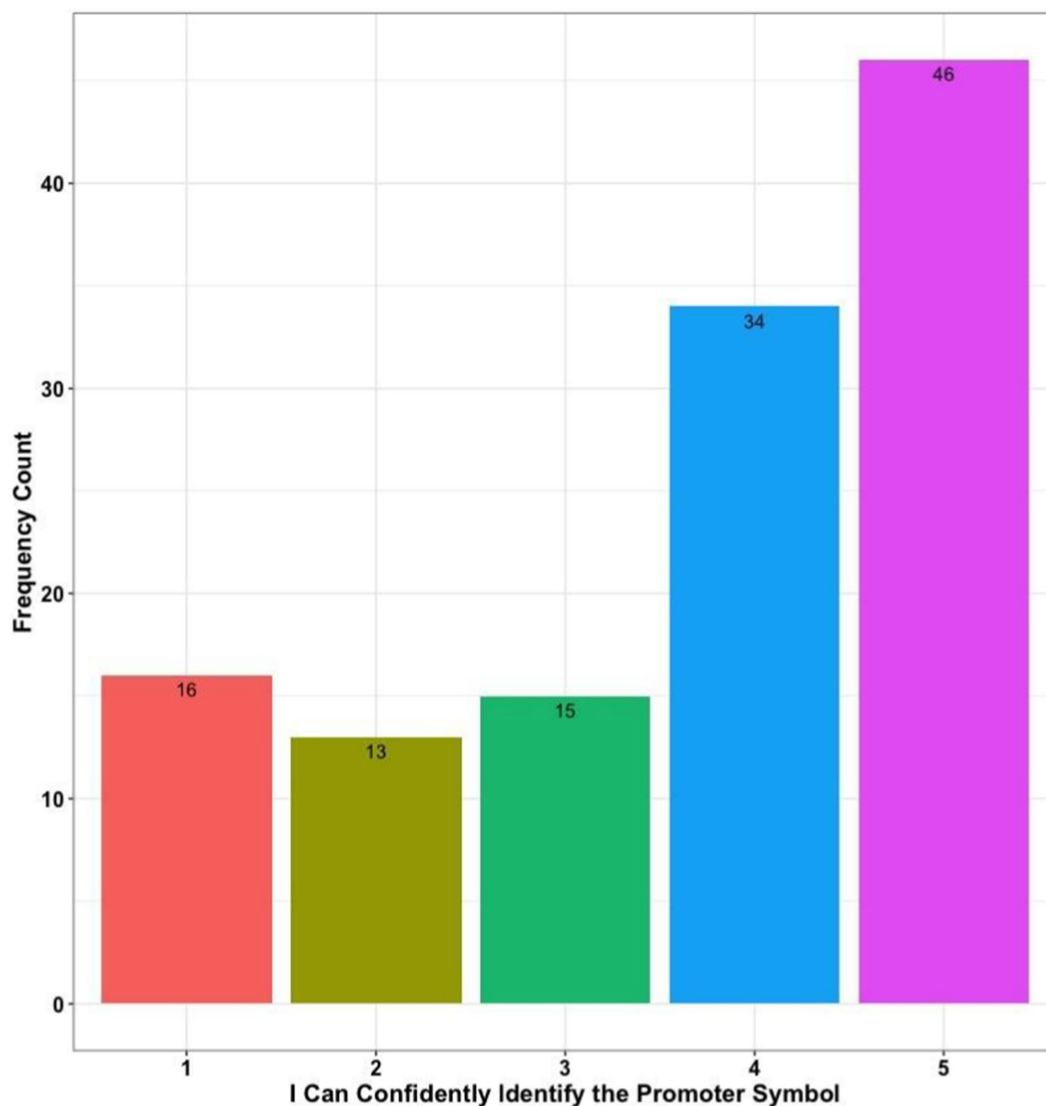
FIGURE 11
Confidence identifying promoter symbol. The item asked if they strongly agree (5) to strongly disagree (1) with the following statement "I can confidently identify the promoter symbol."

Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. LL: Data curation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing. NK: Funding acquisition, Methodology, Project administration, Writing – original draft, Writing – review & editing. CF: Data curation, Project administration, Writing – original draft, Writing – review & editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
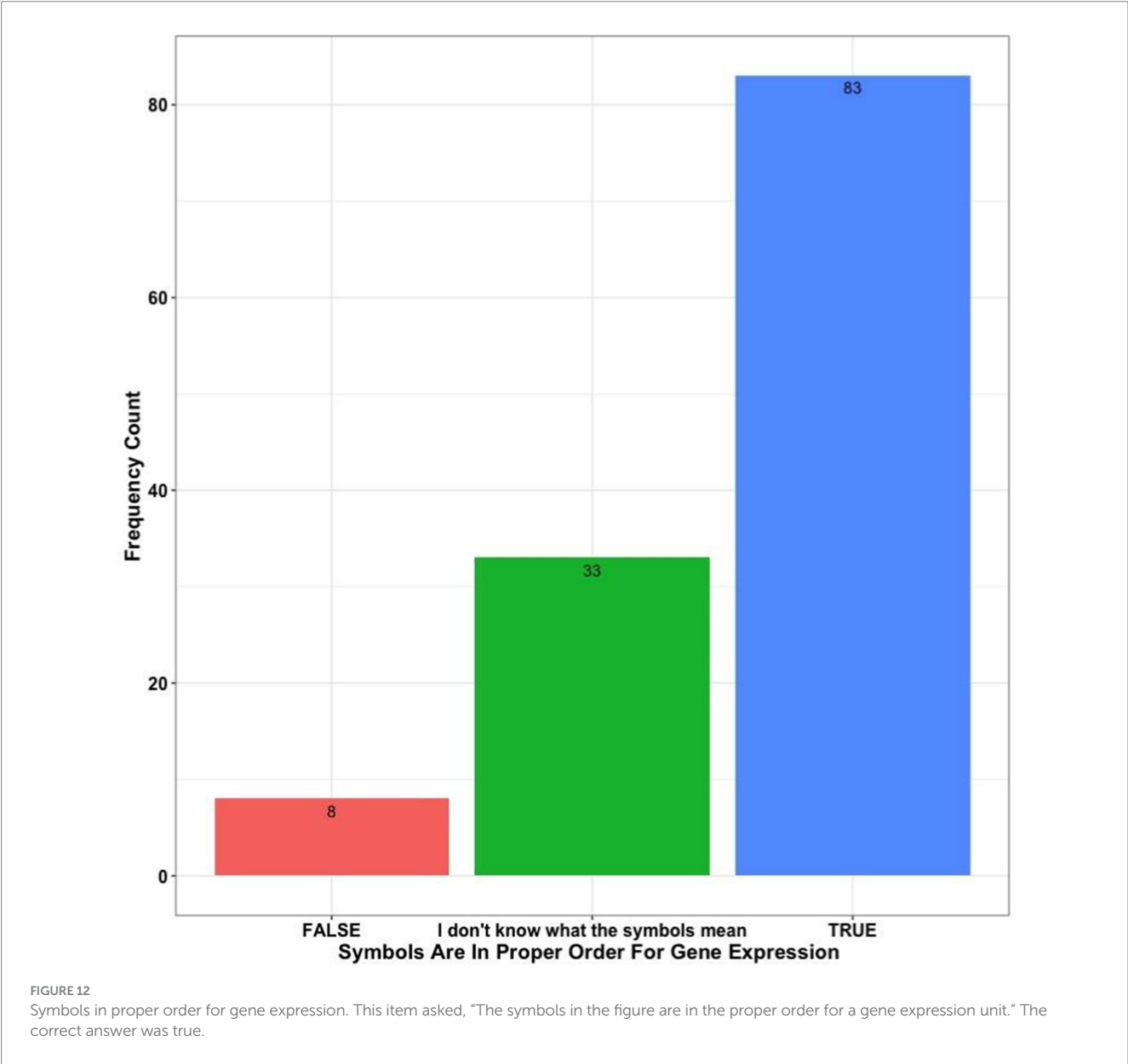
FIGURE 12
Symbols in proper order for gene expression. This item asked, "The symbols in the figure are in the proper order for a gene expression unit." The correct answer was true.

TABLE 8 Synthetic biology (SB) definition rubric.

0	1	2	3
"Do not know"	Poor (wrong or so general as to be a description of any life science)	Emerging (includes idea of design and/or engineering but not target, reason, or approach)	Sophisticated
	That appears to be the process of recreating biological processes and systems from synthetic materials.	Engineering life	SB is a field of science that involves redesigning organisms for useful purposes by engineering them to have new abilities.
	Art	Science of artificial man-made life	Synthetically creating or recreating an organism using pieces of other organisms to solve a problem or create a product.
	Medicine	It is using things like enzymes, proteins, etc., to redesign existing biological systems.	Manipulating parts of gene expression units and combining them w/other ideas to solve a problem.
			Genetically engineering an organism to get it to do what you want.
			SB is idea of creating an engineered organism to accomplish something /to solve a problem.
			It is when biology or science is used to create something that can benefit humans.

TABLE 9 Synthetic biology example rubric.

0	1	2	3
"Do not know"	Poor (incomplete or incorrect)	Emerging (ideas but not connected to process or current applications)	Sophisticated (process, cell or molecular agent, application requires combination that nature would not provide)
	The creation of DNA or genetic	Crisper	An example is redesigning *E. coli* at a transcriptional level so that it will make its media smell like bananas.
	Prosthetics	Production of drugs from cells.	An example could be adding a protein to a cell to make it more useful, like making a color changeable bandages that can react to the pH of our skin wounds.
	Heart monitors	Vaccines.	COVID vaccine (Made using a novel approach that needed synthetic biology)
		Genetically modified food	

TABLE 10 Interrater reliability of responses on synthetic biology content knowledge.

	Definition		Example	
	Pre	Post	Pre	Post
Nominal agreement	100%	92%	96%	88%
Kappa	1.00***	0.85***	0.95***	0.83***

This agreement is based on two expert raters and 25 of 124 responses. Definition = Item asked "What is Synthetic Biology?"; Example = Item asked "Give an Example of Synthetic Biology."
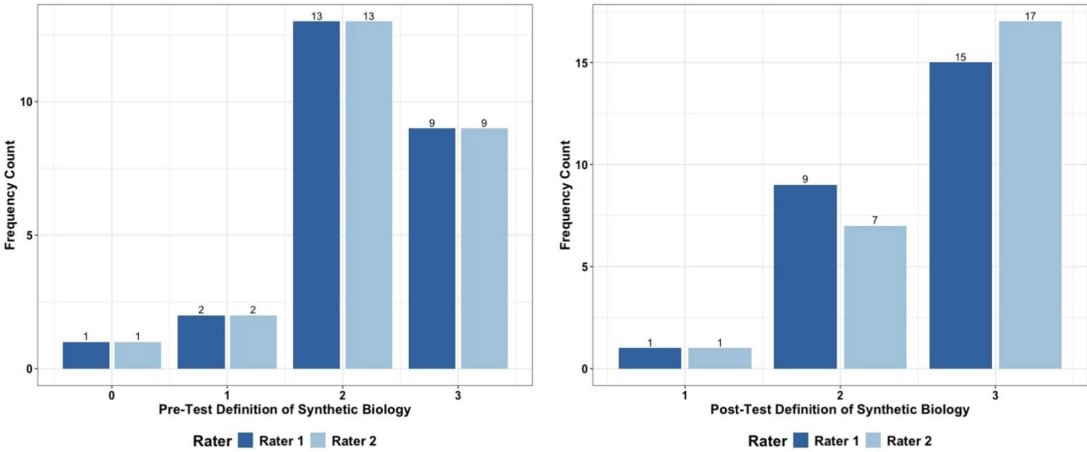*indicates $p <0.05$, **indicates $p < 0.01$; ***indicates $p <0.001$.

FIGURE 13
Pre-post test ratings of student definitions of synthetic biology ($n = 25$). This was a random sampling of 25 students from the 124 responses rated by two expert raters.
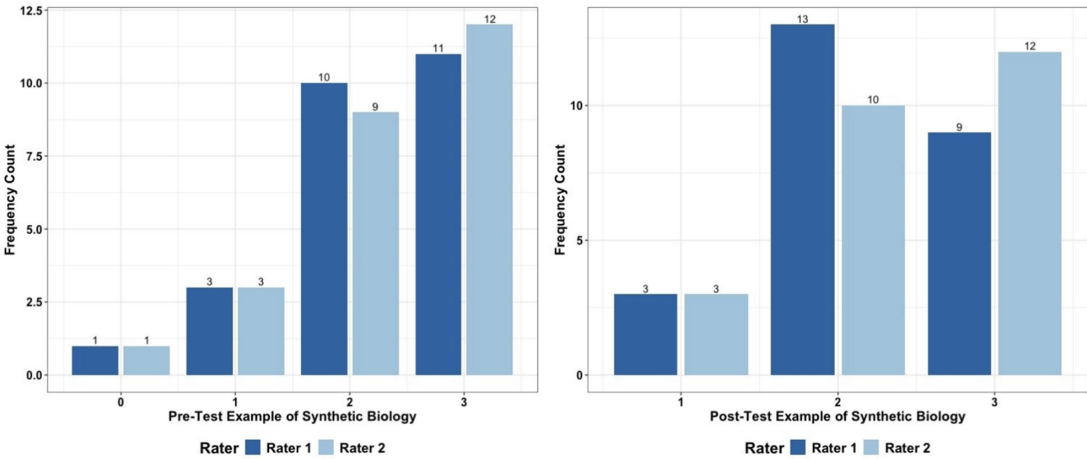

FIGURE 14
Pre-post test ratings of student examples of synthetic biology ($n = 25$). This was a random sampling of 25 students from the 124 responses rated by two expert raters.

References

Allen, J. P., Kuperminc, G., Philliber, S., and Herre, K. (1994). Programmatic prevention of adolescent problem behaviors: the role of autonomy, relatedness, and volunteer service in the teen outreach program. *Am. J. Community Psychol.* 22, 595–615. doi: 10.1007/BF02506896

Aschbacher, P. R., Li, E., and Roth, E. J. (2010). Is science me? High school students' identities, participation and aspirations in science, engineering, and medicine. *J. Res. Sci. Teach.* 47, 564–582. doi: 10.1002/tea.20353

Barton, A. C., Kang, H., Tan, E., O'Neill, T. B., Bautista-Guerra, J., and Brecklin, C. (2013). Crafting a future in science: tracing middle school girls' identity work over time and space. *Am. Educ. Res. J.* 50, 37–75. doi: 10.3102/0002831212458142

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Brickhouse, N. W. (2001). Embodying science: a feminist perspective on learning. *J. Res. Sci. Teach.* 38, 282–295. doi: 10.1002/1098-2736(200103)38:3<282::AID-TEA1006>3.0.CO;2-0

Burke, P. J., and Stets, J. E. (2009). *Identity theory*. Oxford, England: Oxford University Press.

Carlone, H. B., and Johnson, A. (2007). Understanding the science experiences of successful women of color: science identity as an analytic lens. *J. Res. Sci. Teach.* 44, 1187–1218. doi: 10.1002/tea.20237

Castro-Schilo, L., and Grimm, K. J. (2018). Using residualized change versus difference scores for longitudinal research. *J. Soc. Pers. Relat.* 35, 32–58. doi: 10.1177/0265407517718387

Chemers, M. M., Zurbriggen, E. L., Syed, M., Goza, B. K., and Bearman, S. (2011). The role of efficacy and identity in science career commitment among underrepresented minority students. *J. Soc. Issues* 67, 469–491. doi: 10.1111/j.1540-4560.2011.01710.x

Chen, S., Binning, K. R., Manke, K. J., Brady, S. T., McGreevy, E. M., Betancur, L., et al. (2021). Am I a science person? A strong science identity bolsters minority students' sense of belonging and performance in college. *Personal. Soc. Psychol. Bull.* 47, 593–606. doi: 10.1177/0146167220936480

Chin, C., and Osborne, J. (2008). Students' questions: a potential resource for teaching and learning science. *Stud. Sci. Educ.* 44, 1–39. doi: 10.1080/03057260701828101

Chung, J., Cannady, M. A., Schunn, C., Dorph, R., and Vincent-Ruz, P. (2016). *Measures technical brief: competency beliefs in science* Available at: http://www.activationlab.org/wp-content/uploads/2016/02/CompetencyBeliefs-Report-3.2-20160331.pdf (Accessed August 30, 2024).

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 20, 37–46. doi: 10.1177/001316446002000104

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2nd Edn. Hillsdale, New Jersey, USA: Lawrence Erlbaum Associates.

Cumming, G. (2014). The new statistics: Why and how. *Psychol. Sci.* 25, 7–29. doi: 10.1177/0956797613504966

Deemer, E. D., Ogas, J. P., Barr, A. C., Bowdon, R. D., Hall, M. C., Paula, S., et al. (2022). Scientific research identity development need not wait until college: examining the motivational impact of a pre-college authentic research experience. *Res. Sci. Educ.* 52, 1481–1496. doi: 10.1007/s11165-021-09994-6

Farmus, L., Aprin-Cribbie, C. A., and Cribbie, R. A. (2019). Continuous predictors of pretest-posttest change: highlighting the impact of the regression artifact. *Front. Appl. Math.* 4:64. doi: 10.3389/fams.2018.00064

Graham, M. J., Frederick, J., Byars-Winston, A., Hunter, A. B., and Handelsman, J. (2013). Science education. Increasing persistence of college students in STEM. *Science* 341, 1455–1456. doi: 10.1126/science.1240487

iGEM Foundation (2024). *Introduction to participation*. Available at: https://competition.igem.org/participation/introduction (Accessed August 17, 2024).

Jennings, M. A., and Cribbie, R. A. (2016). Comparing pre-post change across groups: guidelines for choosing between difference scores, ANCOVA, and residual change scores. *J. Data Sci.* 14, 205–230. doi: 10.6339/JDS.201604_14(2).0002

Kang, J. (2022). Spanking and children's early academic skills: strengthening causal estimates. *Early Child. Res. Q.* 61, 47–57. doi: 10.1016/j.ecresq.2022.05.005

Kim, A. Y., Sinatra, G. M., and Seyranian, V. (2018). Developing a STEM identity among young women: a social identity perspective. *Rev. Educ. Res.* 88, 589–625. doi: 10.3102/0034654318779957

Kisbu-Sakarya, Y., MacKinnon, D. P., and Aiken, L. S. (2013). A Monte Carlo comparison study of the power of the analysis of covariance, simple difference, and residual change scores in testing two-wave data. *Educ. Psychol. Meas.* 73, 47–62. doi: 10.1177/0013164412450574

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Krajcik, J. S., and Shin, N. (2014). "Project-based learning" in *The Cambridge handbook of the learning sciences*. ed. R. K. Sawyer. *2nd* ed (Cambridge, England: Cambridge University Press), 275–297.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32, 465–491. doi: 10.3102/00028312032003465

Le, H., Robbins, S. B., and Westrick, P. (2014). Predicting student enrollment and persistence in college STEM fields using an expanded P-E fit framework: a large-scale multilevel study. *J. Appl. Psychol.* 99, 915–947. doi: 10.1037/a0035998

Lopatto, D. (2009). *Science in solution: The impact of undergraduate research on student learning*. Council on Undergraduate Research.

Mangu, D. M., Lee, A. R., Middleton, J., and Nelson, J. K. (2015). "Motivational factors predicting STEM and engineering career intentions for high school students" in *2015 IEEE Frontiers in Education Conference: Launching a New Vision in Engineering Education, FIE 2015 - Proceedings* (El Paso, TX: Institute of Electrical and Electronics Engineers Inc).

Marten, G. G. (2022). *Systems thinking: Managing chaos and complexity: A platform for designing business architecture*. Springer.

Mergendoller, J. R., and Thomas, J. W. (2013). *Managing project-based learning: Principles from the field*. Novato, California, USA: Buck Institute for Education.

Metcalf, H. (2010). Stuck in the pipeline: a critical review of STEM workforce literature. *Interactions* 6, 1–21. doi: 10.5070/D462000681

Moore, R., Jensen, M., Haverkamp, C., and Gaston, S. (2011). Science identity development in undergraduates: Rethinking goals and assessment strategies. *CBE Life Sci. Educ.* 10, 321–330. doi: 10.1187/cbe.11-01-0001

Moote, J., Williams, C., and Sproule, J. (2020). The impact of project-based learning on student self-efficacy in science: a longitudinal study. *Int. J. Sci. Educ.* 42, 1528–1544. doi: 10.1080/09500693.2020.1784920

National Science Board and National Science Foundation (2021). *Elementary and secondary STEM education: science and engineering indicators 2022*. NSB-2021-1. Alexandria, VA Available at: https://ncses.nsf.gov/pubs/nsb20211/ (Accessed August 30, 2024).

NISE Network (2024). *Building with biology*. Available at: https://www.nisenet.org/building-with-biology (Accessed August 30, 2024).

R Core Team (2023). *R: a language and environment for statistical computing* [Version 4.3.1]. Vienna, Austria: R Foundation for Statistical Computing.

Renninger, K. A. (2009). Interest and identity development in instruction: an inductive model. *Educ. Psychol.* 44, 105–118. doi: 10.1080/00461520902832392

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *J. Stat. Softw.* 48, 1–36. doi: 10.18637/jss.v048.i02

Shanahan, M. C., and Nieswandt, M. (2009). Creative activities and their influence on identification in science: three case studies. *J. Elem. Sci. Educ.* 21, 63–79. doi: 10.1007/BF03174723

Tabachnick, B. G., and Fidell, L. S. (2001) *Using multivariate statistics* (4th Ed.). Allyn and Bacon.

U.S. Department of Education (2023). *U.S. Department of Education Launches New Initiative to Enhance STEM education for all students*. Available at: https://www.ed.gov/news/press-releases/us-department-education-launches-new-initiative-enhance-stem-education-all-students (Accessed August 13, 2024).

Vincent-Ruz, P., and Schunn, C. D. (2018). The nature of science identity and its role as the driver of student choices. *Int. J. STEM Educ.tion* 5. doi: 10.1186/s40594-018-0140-5

Walker, M. (2021). Exploring middle school students' understanding of synthetic biology. *J. Sci. Educ. Technol.* 30, 428–445. doi: 10.1007/s10956-021-09906-5

Wang, X. (2013). Why students choose STEM majors: motivation, high school learning, and postsecondary context of support. *Am. Educ. Res. J.* 50, 1081–1121. doi: 10.3102/0002831213488622

White House (2022). *FACT SHEET: Biden-Harris administration launches national effort to support student success*. Available at: https://www.whitehouse.gov/briefing-room/statements-releases/2022/07/05/fact-sheet-biden-harris-administration-launches-national-effort-to-support-student-success/ (Accessed August 30, 2024).