2023

Polar Literacy: A Model for Youth Engagement and Learning
Summative Evaluation Report on Youth Outcomes

Prepared by the Utah Education Policy Center
April 2023

THE UNIVERSITY OF UTAH
UTAH EDUCATION POLICY CENTER

THE UNIVERSITY OF UTAH
UTAH EDUCATION POLICY CENTER

Bridging Research, Policy, and Practice

The Utah Education Policy Center (UEPC) is a an independent, non-partisan, not-for-profit research-based center at the University of Utah founded in the Department of Educational Leadership and Policy in 1990 and administered through the College of Education since 2007. The UEPC mission is to bridge research, policy, and practice in public schools and higher education to increase educational equity, excellence, access, and opportunities for all children and adults.

The UEPC informs, influences, and improves the quality of educational policies, practices, and leadership through research, evaluation, and technical assistance. Through our research, evaluation, and technical assistance, we are committed to supporting the understanding of whether educational policies, programs, and practices are being implemented as intended, whether they are effective and impactful, and how they may be improved and scaled-up, and become sustainable.

Please visit our website for more information about the UEPC.
http://uepc.utah.edu

Andrea K. Rorrer, Ph.D., Director
Phone: 801-581-4207
andrea.rorrer@utah.edu

Cori Groth, Ph.D., Associate Director
Phone: 801-581-4207
cori.groth@utah.edu

Ellen Altermatt, Ph.D., Assistant Director for Research and Evaluation
ellen.altermatt@utah.edu

Follow us on Twitter: @UtahUEPC

.

Acknowledgements

The Utah Education Policy Center (UEPC) thanks Janice McDonnell and Kasey Walsh, Rutgers University, and Jason Cervenec and Karina Peggau, The Ohio State University, for their contributions to this report.

Citation: Altermatt, E.R. & Rorrer, A. K. (2023). *Polar Literacy: Summative Evaluation Report on Youth Outcomes.* Salt Lake City, UT: Utah Education Policy Center

Table of Contents

Acknowledgements 3
Table of Contents 4
List of Tables 6
List of Figures 7
Overview 8
Background 8
Goals and Proposed Outcomes 8
Summary of Methods and Key Findings 9
Methods 12
Participants 12
Measures 14
Results 16
Outcome 1. Participating Youth will Report High Levels of Engagement with Science Activities 16
Outcome 2. Participating Youth will Report Increases in Science Identity 18
Outcome 3. Participating Youth will Report Increases in Fascination with Science 20
Outcome 4. Participating Youth will Report Increases in Polar Knowledge 22
Outcome 5. Participating Youth will Report Increases in Confidence in Working with Data 24
Conclusions 28
Recommendations 29
References 31
Addendum 1: Analyses of Data from Winter 2023 Implementation 32
Overview 32
Sample 32
Measures 32
Findings 33
Science Identity 33
Fascination with Science 34
Polar Knowledge 35
Conclusions 36
Addendum 2: Preliminary Analyses Examining the Impact of Specific Implementation Strategies on Youth Outcomes 38
The Impact of Implementation Strategies on Increases in Polar knowledge 38

The Impact of Implementation Strategies on Increases in Science Identity 39

List of Tables

Table 1. Hypothesized Youth Outcomes 9
Table 2. Participation and Demographic Data 12
Table 3. Youths' Reponses to Open-Ended Polar Knowledge Assessment Items 23

List of Figures

Figure 1. Polar Explorer Adventures Homepage 8
Figure 2. Engagement in Science Learning Activity Scores by Program 16
Figure 3. Engagement in Science Learning Activity Scores by Gender and URM Status 17
Figure 4. Science Identity Scores by Time 18
Figure 5. Science Identity Scores by Time and URM Status 19
Figure 6. Science Identity Scores by Time and Initial Science Identity Score 19
Figure 7. Fascination with Science Scores by Time 20
Figure 8. Fascination with Science Scores by Time and Item 20
Figure 9. Fascination with Science Scores by Time and Initial Fascination in Science Scores 21
Figure 10. Polar Knowledge Scores by Time 22
Figure 11. Polar Knowledge Scores by Time and Initial Polar Knowledge Scores 22
Figure 12. Percent of Youth Who Responded "Don't Know" to Polar Knowledge Questions by Time 23
Figure 13. Data Skills Scores by Time 25
Figure 14. Data Skills Scores by Time and Initial Data Skills Scores 25
Figure 15. Confidence in Data Skills by Time 26
Figure 16. Confidence in Data Skills Scores by Time and Initial Confidence 26
Figure 17. Sample "Data Jam" Project from *Data to the Rescue: Penguins Need Our Help!* 27
Figure 18. Science Identity Scores by Time 33
Figure 19. Science Identity Scores by Time and Initial Science Identity Score 34
Figure 20. Fascination with Science Scores by Time 34
Figure 21. Fascination with Science Scores by Time and Initial Fascination with Science Scores 35
Figure 22. Polar Knowledge Scores by Time 35
Figure 23. Polar Knowledge Scores by Time and Initial Polar Knowledge Scores 36
Figure 24. Comparison of Change in Polar Knowledge Scores in Summer 2021 to Fall 2022 Implementations vs. Winter 2023 Implementations 37
Figure 25. Comparison of Change in Polar Knowledge Scores by Presence of a Polar Scientist 38
Figure 26. Comparison of Change in Science Identity Scores by Degree of Choice 39

Overview

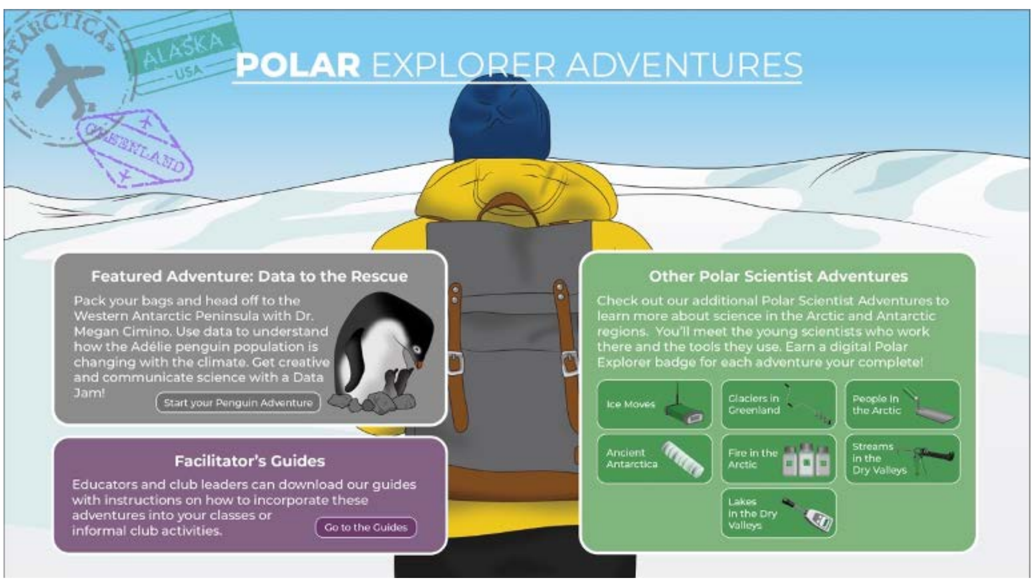
Background

In 2019, the National Science Foundation awarded Janice McDonnell (PI, Rutgers University, Award #1906897) and Jason Cervenec (PI, The Ohio State University, Award #1906929) an Advancing Informal STEM Learning (AISL) grant for a project entitled **Polar Literacy: A Model for Youth Engagement and Learning**. The project was designed to "develop affordable and replicable ways of bringing polar education to informal learning environments, extend our understanding of how polar education initiatives can be delivered to youth with maximum effect, and design a professional development model to improve the capacity for polar region researchers to craft meaningful broader impact activities." The impact of the Polar Literacy project on researchers working in polar regions is addressed in reports provided to Polar Literacy project leaders by the Science Education Resource Center (SERC) at Carleton College. The current report provides a summative assessment of the impact of the Polar Literacy project on proposed outcomes for participating youth.

Goals and Proposed Outcomes

A key goal of the Polar Literacy project was to develop educational materials – including learning kits and videos – that address Polar Literacy Principles[1] and that engage youth in participatory learning opportunities in informal education settings. The expectation is that educational materials will continue to be disseminated and used after the four-year grant period in both online and in person learning environments. Toward this end, project staff are working to ensure that Polar Literacy activities are available online (see Figure 1).

Figure 1. Polar Explorer Adventures Homepage



Among the Polar Literacy resources to be made widely available is *Data to the Rescue: Penguins Need Our Help!* In this hybrid learning activity, youth and an adult facilitator are introduced to polar scientists participating in the Long-term Ecological Research Program (LTER) at Palmer Station on

[1] https://polar-ice.org/polar-literacy-initiative/

the Western Antarctic Peninsula. Through a series of guided activities, participants are invited to explore how the Antarctic ecosystem is transforming due to climate-related changes including melting ice, increased ocean temperatures, and changing animal populations. For example, participants are introduced to a biological oceanographer who studies penguins in Antarctica, gain practice in reading graphs, interact with an NSF-funded data visualization tool called CODAP (Common Online Data Analysis Platform), use CODAP to explore how the populations of three different types of penguins has changed over time, and share what they have learned about changing penguin populations through a creative project called a Data Jam. Project leaders have received supplemental funding from NSF to work with a graphic design team, a videographer, and a case study research team to finalize *Data to the Rescue* materials. These materials will be reviewed by the national 4-H Council for potential mass distribution.

Project leaders hypothesize that engaging with these and other Polar Literacy activities will result in the five positive youth outcomes summarized in Table 1.

Table 1. Hypothesized Youth Outcomes

Outcome	Description
1	Participating youth will report high levels of **engagement in science activities**.
2	Participating youth will report increases in **science identity**.
3	Participating youth will report increases in **fascination with science**.
4	Participating youth will report increases in **polar knowledge**.
5	Participating youth will report increases in **confidence in working with data**.

To permit preliminary testing of these hypotheses and to inform both the development of curricular materials and the development of instructional methods for implementing these materials to maximize impact, Polar Literacy project staff hosted multiple learning opportunities for youth from Summer 2020 to Fall 2022. These opportunities were offered across a wide range of settings (including in schools, afterschool settings, community centers, and camps) and in a variety of formats (including virtual, in-person, and hybrid).

Summary of Methods and Key Findings

Across 28 sites, 453 youth who participated in Polar Literacy programming assented – and received parental consent – to participate in the evaluation of the program by completing surveys designed to measure their engagement with program activities and changes in their science attitudes and knowledge.[2] Review of program documents, observations of program activities, and analyses of survey data yielded the following key findings. ***Detailed findings are provided in the remainder of this report followed by recommendations for continuous improvement.***

1. Across program sites, participating youth reported high levels of engagement in science activities. Engagement scores were equally high among male and female youth and among youth who did and did not identify as underrepresented minorities (URMs). Youth expressed especially strong engagement when activities allowed them to take the lead in discovery (e.g., by solving mysteries), focused on animals, focused on data, and allowed them to engage in hands-on learning. Observations by external evaluators indicated that

[2] Any opinions, findings, and conclusions or recommendations expressed in this document are those of the authors and do not necessarily reflect the views of the Principal Investigators or the National Science Foundation.

youth engagement was highest when opportunities to participate were anonymous or low-stakes, mechanisms for recognition were available (e.g., positive feedback, badges), youth were given specific roles to play during sessions (e.g., note taker or reporter), and youth had time to actively engage with one another and with project leaders to develop a sense of community.

2 Across program sites, participating youth reported statistically significant increases in science identity. Gains were especially strong among youth who began the program with lower science identity scores. Youth who identified as underrepresented minorities (URMs) reported lower science identity than youth who did not identify as URMs. Importantly, however, youth who identified as URMs experienced gains in science identity from pre-program to post-program that were similar in magnitude to the gains made by youth who were not URMs. These findings are consistent with prior research indicating that science identity can be bolstered among youth by employing strategies that include a) helping students see scientists as people to whom they can relate and b) showcasing scientists who don't match stereotypes (e.g., Basu & Barton, 2007; Schinske et al., 2016). Both strategies were important components of Polar Literacy programming. For example, at several program sites, youth had an opportunity to interact directly with Polar Scientists, many of whom were young, women, or members of a URM group and who took time to share their personal interests and stories.

3 Participating youth did not, as a group, report increases in fascination with science. However, youth with initially lower levels of fascination did show statistically significant gains from pre-program to post-program. These results should be interpreted with some caution given that only two items were included in surveys across program sites and lower ratings on one of these items – "I want to know how to do everything scientists do." – may not be problematic if lower ratings reflect a growing understanding that scientists do many different things and that some activities are more fascinating than others. Consistent with this interpretation, two common refrains in youths' responses to open-ended questions regarding program impact were that their interest in science had grown (e.g., "I find science quite interesting now."), but that they also had new appreciation for how difficult the work that polar scientists do can be (e.g., "It made me realize the hard work put in.")

4 Participating youth did not, as a group, show increases in polar knowledge on a multiple-choice assessment. However, youth with initially lower levels of polar knowledge did show statistically significant gains from pre-program to post-program on this assessment. Moreover, gains were evident in students' responses to open-ended questions tapping their polar knowledge. On pretest assessments, nearly 30% of youth indicated that they "did not know" how polar regions have been affected by climate change or what tools researchers in polar regions used. On posttest assessments, fewer than 10% of youth indicated that they did not know. Importantly, youths' post-survey responses reflected an understanding of key Polar Literacy Principles. In response to open-ended questions tapping their perceptions of what they learned about polar science or the work that polar scientists do, youth were especially likely to describe changes in three areas: a) a greater understanding of the characteristics of the polar regions, b) a greater understanding of and appreciation for the plight of polar animals, and c) a greater understanding of and appreciation for the work that polar scientists do.

5 Across program sites, participating youth reported statistically significant increases in confidence in working with data. Gains were especially strong among youth who began the program with lower confidence. These findings are consistent with evidence that data skills

can be bolstered in youth when instruction includes opportunities for youth to create and interact with high-quality visualizations (Krumhansl, Peach, Oster, Busey, & Baker, 2012), uses real-world datasets (Erwin, 2015), and is problem-based and student-directed (English & Watson, 2015). All three strategies were important components of Polar Literacy programming.

Methods

Participants

Program staff reported that more than 600 youth participated in Polar Literacy programming from Summer 2020 to Fall 2022 at 28 sites.[3] Table 2 provides an overview of participation data and demographic data for each program site through Fall 2022. All analyses of survey data are based on the 453 youth participants who assented to participate in evaluation activities and who also had parental consent during this time period. Importantly, in an Addendum to this report, we present findings from data collected in Winter 2023 that were not yet available at the time that data collected through Fall 2022 were analyzed.

Participation Data. Participation data was provided by Polar Literacy program staff. As shown, the number of youth participants ranged from 8 to 166 across program sites. Most youth were middle-school-aged and the number of youth participants increased markedly over time.

Demographic Data. Demographic data was derived from registration information collected by Polar Literacy program staff and is included *only* for youth participants who received parental consent and provided child assent (n = 453). Of these, 186 (41.1%) had parents who provided information on race/ethnicity and 293 (64.7%) had parents who provided information on gender. Among youth whose parents provided demographic data, the majority (61.3%) had unrepresented minority (URM) status[4] and a slight majority identified as female (52.9%).

Table 2. Participation and Demographic Data

Season	Program Site Descriptions (Format, Location)	# of youth participants	# of youth with consent	Grade/Age Range	% Female	% URM
Summer 2020	Polar Week #1 (Virtual, Ohio)	19	19	6th – 8th	53%	6%
	Polar Week #2 (Virtual, Newark, NJ)	13	12	4th – 9th	42%	92%
	Polar Week #3 (Virtual, Trenton, NJ)	23	12	4th – 8th	67%	100%
Fall 2020	Polar Research Expedition (Virtual, Colorado)	11	9	6th – 8th	22%	22%
	Polar Mystery (Virtual, Ohio)	14	14	6th – 8th	57%	NA
Winter 2021	Arctic Mystery (Virtual, Ohio)	18	17	6th – 8th	35%	13%
	Succeed2gether (Virtual, Montclair, NJ)	17	15	5th – 8th	47%	NA

[3] Program staff reported 115 participants in Year 1 (i.e., from Summer 2020 to Winter 2021) and 461 participants in Years 2 (i.e., from Summer 2021 to Winter 2022), representing a four-fold increase in the number of youth served.
[4] Youth were identified as having Underrepresented Minority (URM) status if parents or caregivers indicated on Polar Literacy registration forms that youth were Black or African American, Hispanic or Latino, or American Indian or Alaska Native (National Science Foundation, 2020).

Season	Program Site Descriptions (Format, Location)	# of youth participants	# of youth with consent	Grade/Age Range	% Female	% URM
Summer 2021	4-H from Home (Virtual, multi-state)	45	19	NA	58%	NA
	Paterson Charter (in person, New Jersey)	25	2	NA	0%	NA
	The Franklin Institute (in person, Pennsylvania)	NA	12	14 – 15 yrs	42%	67%
	STEM Ambassadors (Virtual, New Jersey)	43	18	11 – 16 yrs	72%	80%
	Ohio 4-H and HFF (in person, Ohio)	16	9	9 – 14 yrs	NA	NA
	Dodd Street Housing (in person, New Jersey)	15	0	NA	NA	NA
	YMCA (in person, Ohio)	22	7	NA	NA	NA
Fall 2021	Gladden House (in person, Ohio)	30	27	6th – 9th	NA	NA
	Paterson Charter (in person, New Jersey)	53	6	NA	33%	100%
	Asynchronous (Virtual)	NA	5	NA	60%	0%
	Jersey City PS5 (Hybrid, New Jersey)	166	38	NA	42%	59%
	Neptune MS and HS (in person, New Jersey)	90	25	NA	44%	67%
	Ohio Hispanic Coalition (in person, Ohio)	10	12	K-8th	42%	100%
Winter 2022	Camp Oty'Okwa (in person, Ohio)	66	42	7th	NA	NA
	4-H SPIN (Virtual)	8	8	6th – 8th	NA	NA
Summer 2022	Horizon STEM Acad (in person, Ohio)	12	8	9th – 11th	NA	NA
	Jersey City #23 (in person, Jersey City, NJ)	NA	58	2nd – 8th	45%	NA
	Jersey City #34 (in person, Jersey City, NJ)	NA	14	2nd – 9th	57%	NA
	Ohio State Fair (in person, Ohio)	18	18	3rd – 9th	NA	NA
Fall 2022	St. Charles (in person, Ohio)	29	3	8th	NA	NA
	Worthingway (in person, Ohio)	14	7	6th – 8th	NA	NA

Note. "NA" indicates that these data were not available. Ns for # of youth with consent do not sum to 453 as a program site description was missing for 17 youth.

Measures

Youth participants were asked to complete surveys at multiple time points. To test hypotheses about the proposed impact of the Polar Literacy project on student outcomes, the surveys were designed to assess youths' engagement in science learning and changes in youths' science identity, fascination in science, polar knowledge, and data skills.

1. Engagement in Science Learning Activities. Engagement in program activities was measured using eight items from the Engagement in Science Learning Activities Survey.[5] Youth were asked to respond to each item (e.g., "During this activity, I felt excited") on a four-point scale that ranged from 1 (e.g., "NO!") to 4 ("YES!"). Prior to averaging youths' responses across the eight items, four items were reverse scored so that higher scores always indicated higher engagement. The eight items formed a reliable scale with Cronbach's alphas (αs) ranging from .69 to .83. Programs lasting more than three days were not included in reliability analyses because of small *n*s.[6]

2. Science Identity. Across time and program sites, science identity was measured at two or more time points using up to seven items modified from Cole (2012). For this report, analyses focused on the analysis of three items that were consistent across time and program sites. Pre-program survey items were administered at the beginning of programming at each site. Post-program survey items were administered at the end of programming at each site. Youth were asked to respond to each item (e.g., "I am interested in pursuing a career in a scientific field") on a five-point scale that ranged from 1 ("not at all true") to 5 ("really true"). Scores on each item were averaged so that higher scores indicated stronger science identity. The three items formed a reliable scale with Cronbach's alphas (αs) ranging from .78 to .86 across the two time points.

3. Fascination in Science. Across time and program sites, fascination in science was measured at two time points (i.e., at the beginning and end of the program) using up to eight items from the Fascination in Science subscale of the Science Learning Activation Survey (2016).[7] For this report, analyses focused on the analysis of two items that were consistent across time and program sites. Youth were asked to respond to each item (e.g., "I want to know how to do everything that scientists do") on a four-point scale that ranged from 1 (e.g., "NO!") to 4 ("YES!"). Youths' responses were averaged across the two items so that higher scores always indicated greater fascination in science. The two items formed a reliable scale with Cronbach's alphas (αs) ranging from .72 to .83 across the two time points.

4. Polar Knowledge. Beginning in Summer 2021, youth participants were asked to complete a Polar Knowledge assessment at two time points (i.e., at the beginning and end of the program). The assessment includes four items designed by Polar Literacy program staff and modeled on items developed for a national assessment of students' polar knowledge (Pfirman et al., 2021). Each item offers three response options. For the first two items, youth could select more than one response. For these items, each correct response (i.e., selecting a correct response or not selecting an incorrect option) was awarded .5 points while each incorrect selection (i.e., failing to select a correct option or

---

[5] See http://activationlab.org/wp-content/uploads/2018/03/Engagement-Report-3.2-20160803.pdf

[6] Cronbach's alpha (α) is a measure of internal reliability, or the degree to which items hold together as a group. When items hold together as a group, they can be combined (e.g., averaged) to form a scale. A scale is generally considered to be internally reliable if a is equal to or greater than .7. Importantly, a high α does not imply that the items form a unidimensional scale.

[7] See http://activationlab.org/wp-content/uploads/2018/03/Fascination-Report-3.2-20160331.pdf

selecting an incorrect option) was assigned -.5 points. For the third and fourth items, youth could select only one response. For these items, a correct response was awarded 1.5 points and an incorrect response was assigned -1.5 points. Thus, the possible score for each item ranged from -1.5 to 1.5. Scores were summed across items so that higher scores (ranging from -6 to +6) represented stronger polar knowledge.

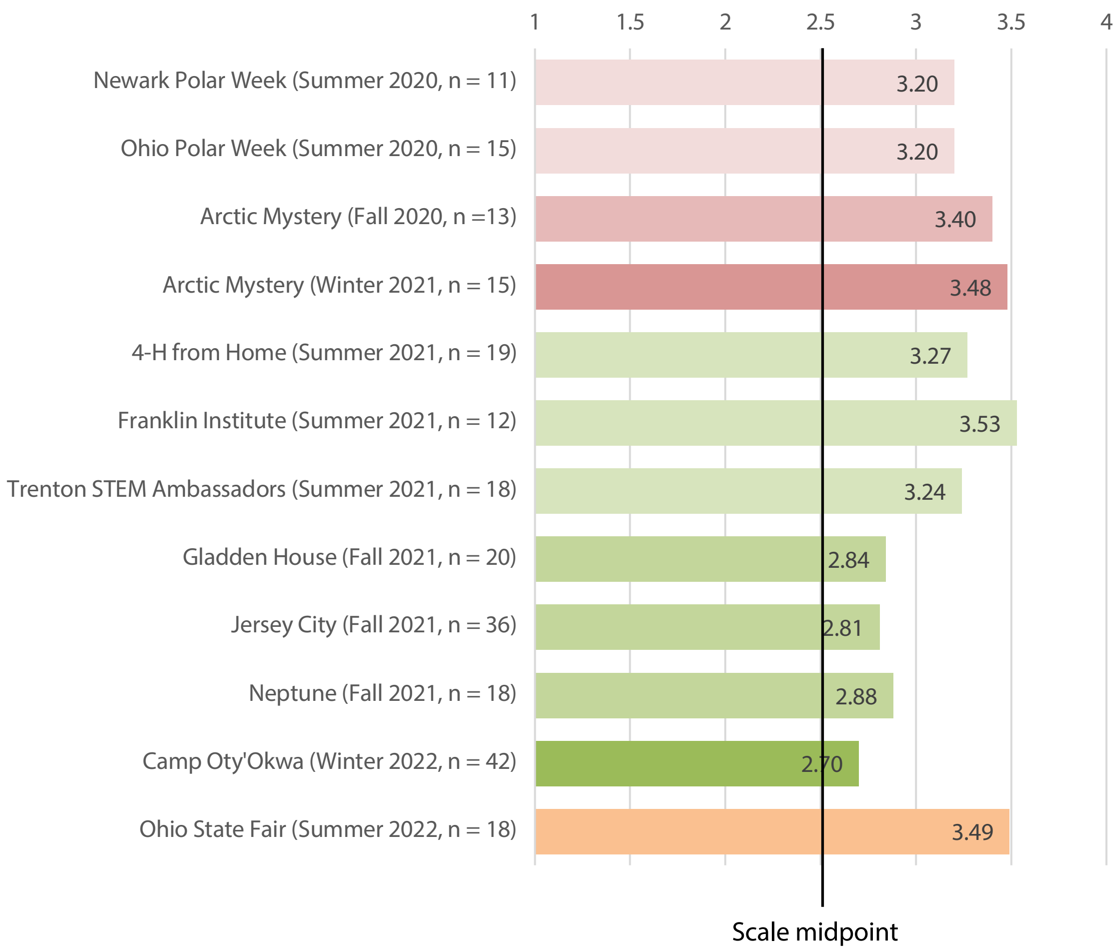
5. **Data Skills.** In Summer 2021, youth participants were asked to complete a data skills self-assessment at two time points (i.e., at the beginning and end of the program). The assessment includes four items designed by Polar Literacy program staff and an associate (Kristin Hunter-Thomson). A correct response was awarded 1 point and an incorrect response was awarded 0 points. Thus, possible scores for the data skills assessment ranged from 0 to 4. This assessment was discontinued after Summer 2021. After Summer 2021, youth participants were asked to rate their confidence in their data skills before and after the program using a single post-survey item on a scale that ranged from 1 (not at all confident) to 5 (very confident). Higher scores represented stronger confidence in data skills. The latter measure is the focus of this report given larger sample sizes.

Results

Outcome 1. Participating Youth will Report High Levels of Engagement with Science Activities

Across programs, 310 youth responded to the Engagement in Science Learning Activities survey, with the number of respondents ranging from 1 to 41 across programs. Engagement scores for all programs with 10 or more respondents are shown in Figure 2, sorted by event (with the most recent events appearing last). As shown, self-reported engagement was high. Scores, averaged across days, were above the midpoint of the four-point scale across all 12 programs. Averaging across days and programs, the mean engagement score was also well above the scale midpoint (mean = 3.10).

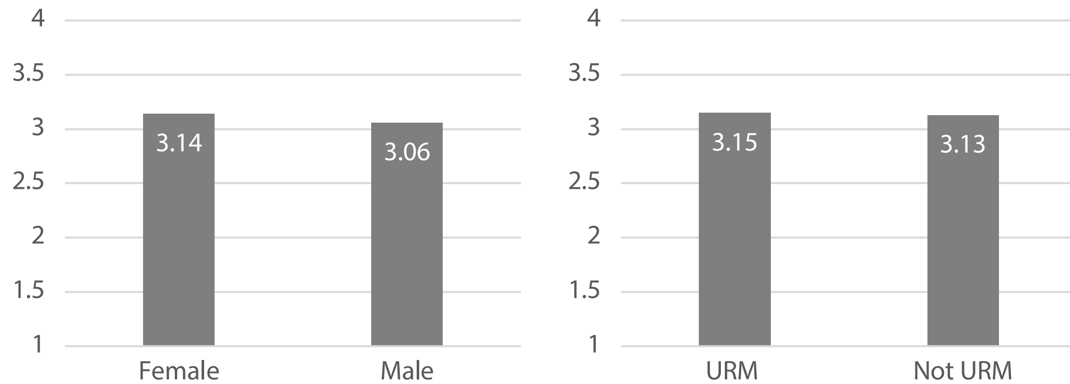
Figure 2. Engagement in Science Learning Activity Scores by Program



Note. Only programs for which the number of respondents was greater than or equal to 10 are listed.

Importantly, results of a two-way analysis of variance (ANOVA)[8] indicated that levels of engagement did not vary either by gender or by URM status. As shown in Figure 3, students who identified as male and female and who identified as URMs and not as URMs had engagement scores that were well above the midpoint of the four-point scale across days and programs.

Figure 3. Engagement in Science Learning Activity Scores by Gender and URM Status



Responses to an open-ended survey item asking students to indicate which program activities they enjoyed most revealed that youth were especially enamored of opportunities that a) allowed them to take the lead in discovery (e.g., "being a detective", "solving the cases"), b) focused on animals (e.g., "learning about penguins and their habitats"), c) focused on data (e.g., "the data jam", "m&m data"), and d) allowed them to engage in hands-on learning (e.g., with "clay," "slime," and "arts and crafts").

Observations by external evaluators indicated that youth engagement was highest when:

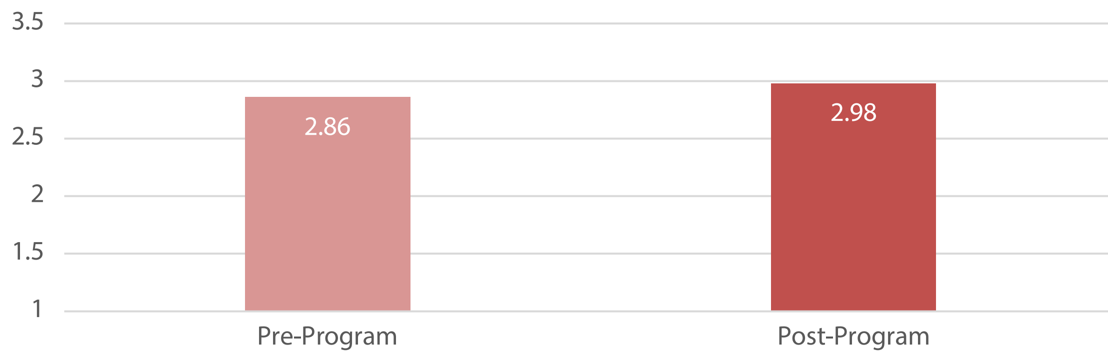
- opportunities for youth to participate were anonymous or low-stakes (e.g., when youth were asked to participate in anonymous polls or to share their responses in Zoom's "chat" feature);
- program leaders allowed sufficient "wait time" after asking questions;
- youth were able to interact in smaller break-out rooms;
- collaborative technology tools were available and functioning as expected (e.g., the annotation feature on Zoom);
- youth were directly invited to engage (e.g., "Juan, what do you think?");
- opportunities for recognition (e.g., positive feedback, badges) were available;
- youth had an opportunity to interact with program materials sent to them via mail before the virtual sessions;
- youth were given specific roles to play during sessions; and
- leaders allowed youth time to actively engage with one another to allow a sense of "community" to form as the program progressed.

[8] Analysis of variance is used to examine the null hypothesis that there is no difference between the mean scores of different groups. The larger the F value, the greater the evidence against the null hypothesis. Here, the null hypothesis is that there is no difference in mean engagement in science learning activity scores across gender or URM status.

Outcome 2. Participating Youth will Report Increases in Science Identity

Across sites, 205 students had both pre-program and post-program science identity scores. As shown in Figure 4, results of a dependent t-test[9] indicated that there was a statistically significant increase in youths' science identity scores from pre-program (mean = 2.86) to post-program (mean = 2.98), $t(204) = 2.35$, $p < .05$.[10] The effect size for the increase was $d = .16$ which, according to Kraft (2020), represents a "medium" effect compared to other educational interventions.[11] The largest difference emerged for the item that read "I am similar to a professional scientist." Here, there was an even stronger, statistically significant increase in youths' science identity scores from pre-program (mean = 2.19) to post-program (mean = 2.47), $t(204) = 4.12$, $p < .001$. The effect size for the increase was $d = .29$ which, according to Kraft (2020), represents a "large" effect compared to other educational interventions.

Figure 4. Science Identity Scores by Time



Repeated measures analyses of variance (ANOVAs) with time as the within-subjects factor and gender and URM status as between-subjects factors indicated there was a main effect for URM status on science identity such that youth who were URMs had lower science identity (estimated marginal mean = 2.62) than youth who were not URMs (estimated marginal mean = 3.17), $F(1, 70) = 6.17$, $p < .05$. Importantly, however, URM status did not interact with Time, indicating that youth who were URMs experienced gains in science identity from pre-program to post-program that were similar in magnitude to the gains made by youth who were not URMs (see Figure 5). Although one might have hoped for even larger gains among youth who were URMs to close the gap in science identity scores by URM status, the finding of similar gains is still encouraging.

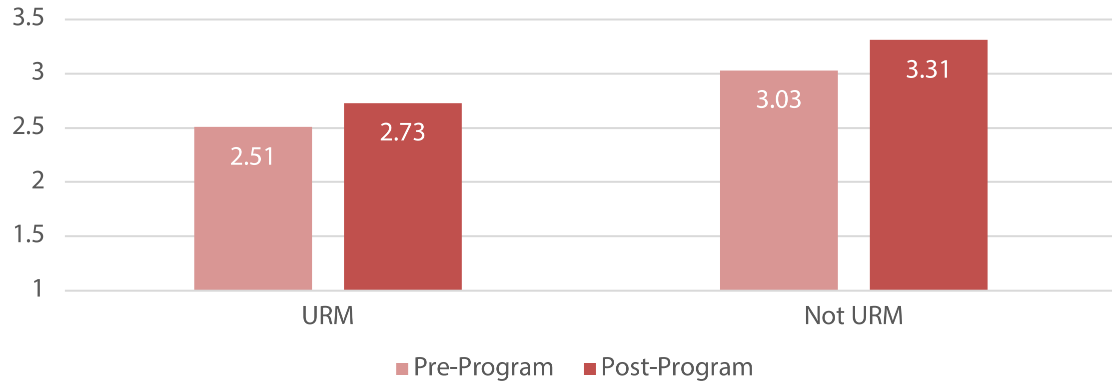
---

[9] A dependent t-test is used to examine differences between scores when participants are measured on the same outcome variable at two time points. The larger the *t* value, the greater the evidence against the null hypothesis. Here, the null hypothesis is that there is no difference between pre-survey and post-survey responses.

[10] A small p value indicates strong evidence against the null hypothesis. When $p < .05$, the chances of finding a difference this large or larger by chance (given that the null hypothesis is true) is less than 5%. Such a result is said to be statistically significant.

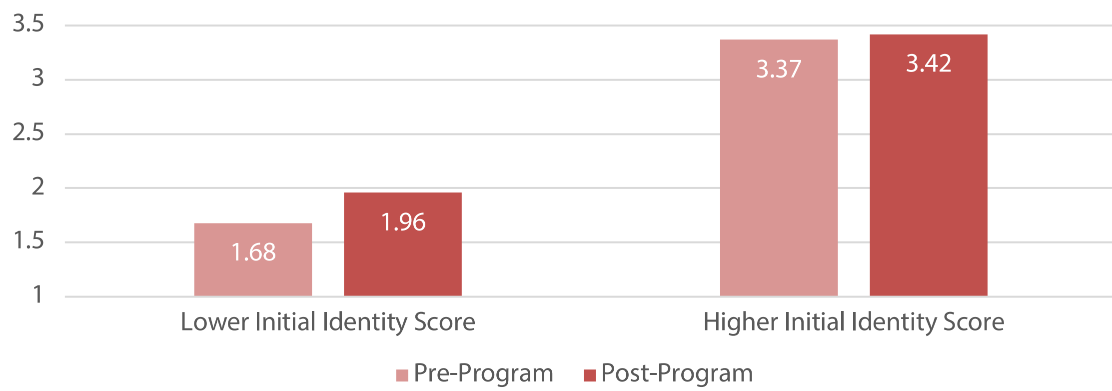
[11] Kraft (2020) proposes that an effect size less than 0.05 is small, 0.05 to less than 0.20 is medium, and 0.20 or greater is large. These proposed benchmarks are based on the distribution of 1,942 effect sizes from 747 Randomized Control Trials (RCTs) evaluating education interventions with standardized test outcomes. Caution should be taken in applying these benchmarks – or any other benchmarks – to the current analyses as debate about the most appropriate interpretation of effect sizes is ongoing (see Bakker, Cai, English, Kaiser, Mesa, & Van Dooren, 2019).

Figure 5. Science Identity Scores by Time and URM Status



To explore whether the impact of Polar Literacy programming might differ for youth with higher versus lower initial science identity scores, the sample was divided into those with pre-program scores lower than 2 (the "lower initial identity score" group; $n = 62$) and those with scores greater than 2 (the "higher initial identity score" group, $n = 143$). As shown in Figure 6, a repeated measures analysis of variance (ANOVA) with time as the within-subjects factor and initial science identity score as the between-subjects factor indicated that there was a statistically significant difference between the two groups, $F(1, 203) = 4.83$, $p < .05$, with the lower initial identity score group showing greater gains than the higher initial identity score group.

Figure 6. Science Identity Scores by Time and Initial Science Identity Score

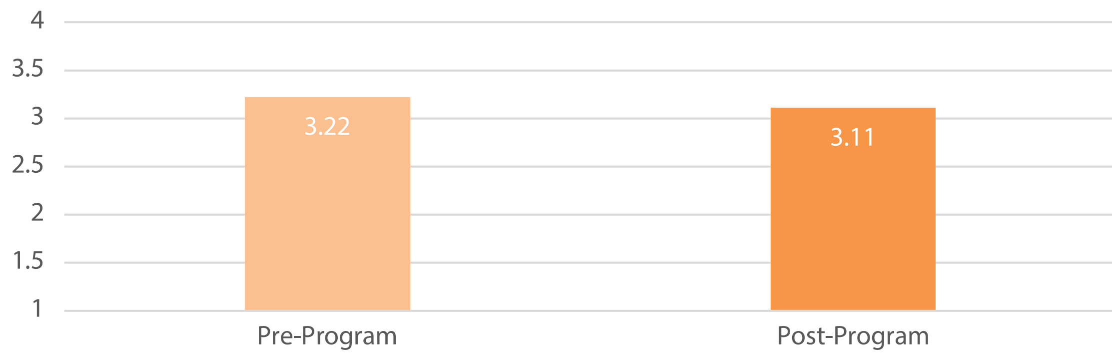


These findings are consistent with prior research indicating that science identity can be bolstered among youth by employing strategies that include a) helping students see scientists as people to whom they can relate and b) showcasing scientists who don't match stereotypes (e.g., Basu & Barton, 2007; Schinske et al., 2016). Both strategies were important components of Polar Literacy programming. For example, at several program sites, youth had an opportunity to interact directly with polar scientists, many of whom were young, women, or members of a URM group and who took time to share their personal interests and stories. At other sites, youth had an opportunity to "meet" these scientists by watching a video.

Outcome 3. Participating Youth will Report Increases in Fascination with Science

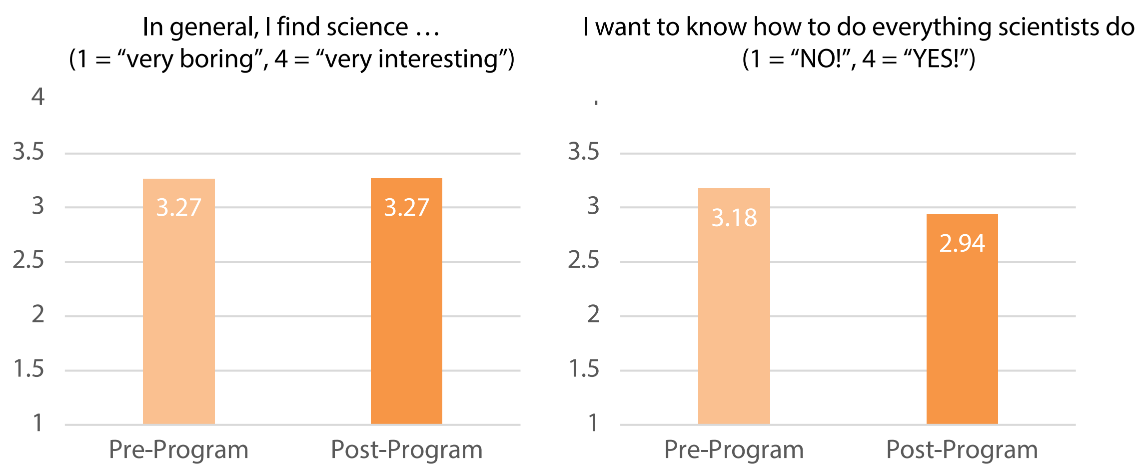
Across sites, 202 students had both pre-program and post-program fascination with science scores. As shown in Figure 7, results of a dependent t-test indicated that there was a statistically significant *de*crease in youths' fascination scores from pre-survey (mean = 3.22) to post-survey (mean = 3.11), $t(201) = 2.73$, $p < .01$.

Figure 7. Fascination with Science Scores by Time



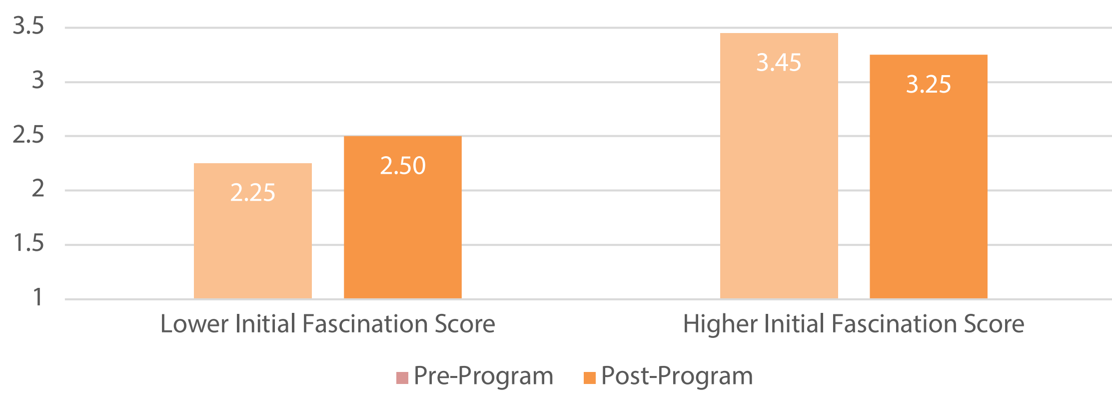
In interpreting this finding, it is important to note that the trend did not differ by gender or URM status, but the trend *did* differ across items and by youth's initial levels of fascination. As shown in Figure 8, there was no change from pre-program to post-program in youths' perceptions of the degree to which science was interesting. At both time points, the mean rating was well above the midpoint of the scale (means = 3.27 on a 4-point scale). There was, however, a statistically significant decrease in the degree to which youth reported wanting to do everything scientists do from pre-program (mean = 3.18) to post-program (mean = 2.94), $t(199) = 3.86$, $p < .001$. In future evaluations, follow-up interviews with youth might be helpful in interpreting this finding. On the one hand, the decline could reflect a true decline in fascination with science. On the other hand, the decline could reflect an enhanced understanding that scientists do *many different things* and a realization among youth that some activities are more fascinating than others.

Figure 8. Fascination with Science Scores by Time and Item



The pattern of findings also differed by youth's initial fascination with science. To explore these differences, the sample was divided into those with pre-program scores lower than or equal to 2.5 (the "lower initial fascination score" group; $n = 38$) and those with scores greater than 2.5 (the "higher initial fascination score" group, $n = 164$). As shown in Figure 9, a repeated measures analysis of variance (ANOVA) with time as the within-subjects factor and initial fascination score as the between-subjects factor indicated that there was a statistically significant difference between the two groups, $F(1, 200) = 19.84$, $p < .001$, with the lower initial fascination score group showing statistically significant gains in fascination, $t(37) = 2.84$, $p < .01$, and the higher initial fascination score group showing statistically significant declines, $t(163) = 4.48$, $p < .001$. Again, these findings should be interpreted with some caution given that only two items were included in surveys across time points and the interpretation of one of these items – as noted above – is ambiguous.

Figure 9. Fascination with Science Scores by Time and Initial Fascination in Science Scores



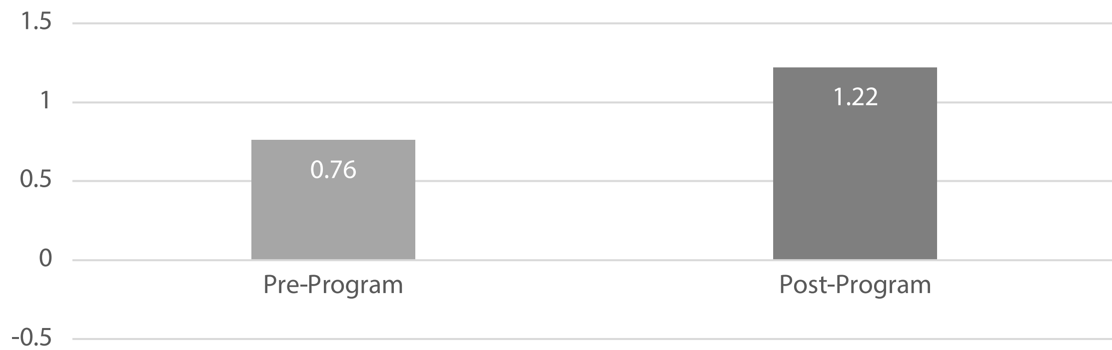
That some youth experienced gains in fascination with science as a result of their participation is supported by students' responses to open-ended items assessing the impact of Polar Literacy programming. A common – albeit not universal – refrain in students' comments was that the program increased their interest in science and their desire to learn more. At the same time, youth indicated that they had new appreciation for how difficult the work that polar scientists do can be. For example:

> *"I find science quite interesting now."*
>
> *"[The program] made me more interested."*
>
> *"[The program] made me want to learn more."*
>
> *"It made me realize the hard work put in."*
>
> *"Polar scientists work very hard."*
>
> *"It helped me understand about polar science that they work really hard."*

Outcome 4. Participating Youth will Report Increases in Polar Knowledge

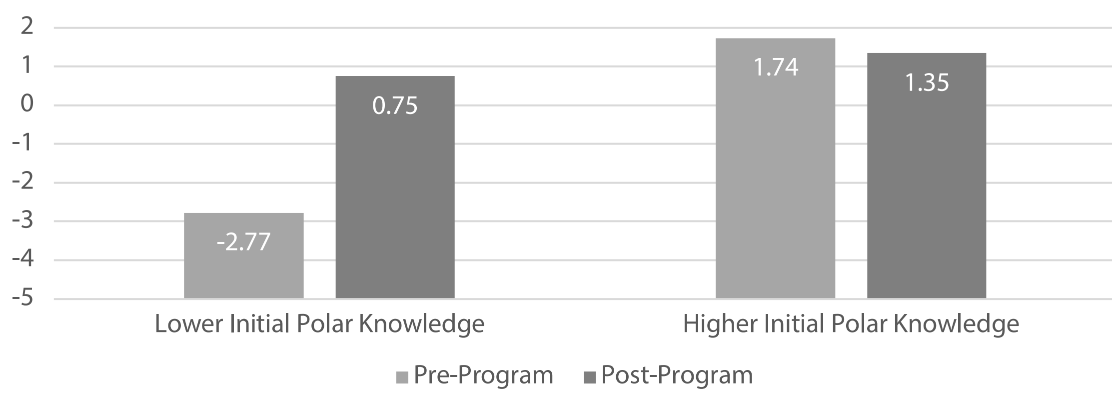
Across sites, 101 students had both pre-program and post-program Polar Knowledge scores. As shown in Figure 10, polar knowledge scores at post-program (mean = 1.22) were higher than those at pre-program (mean = .76).

Figure 10. Polar Knowledge Scores by Time



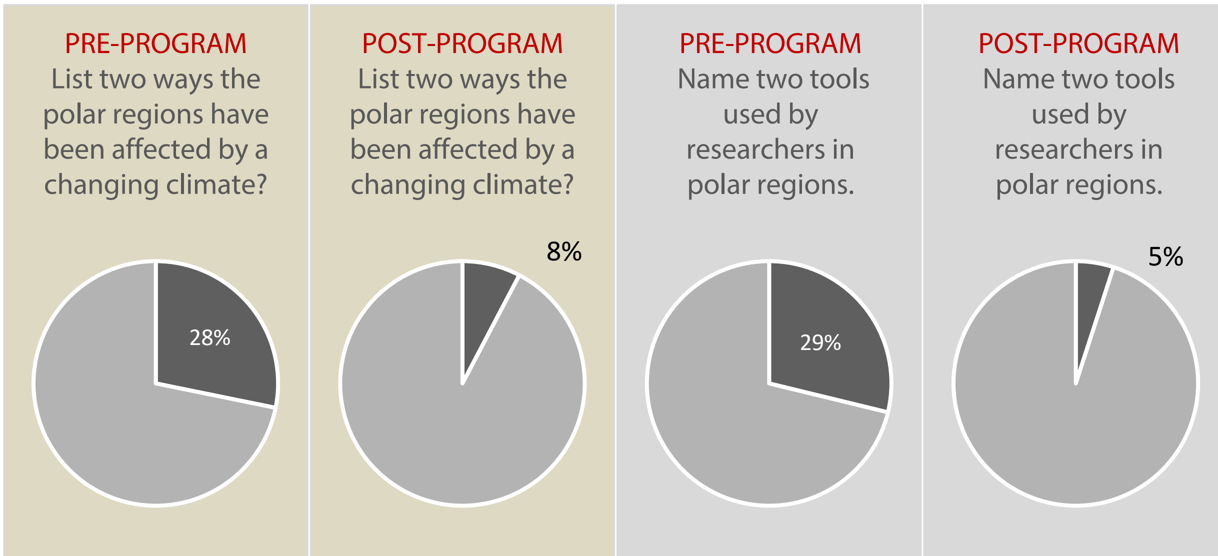
Although results of a dependent *t*-test indicated that this difference was not statistically significant, $t(100) = 1.26$, $p > .05$, and did not differ by student gender or URM status, it is, again, possible that the pattern of findings differs by youth's initial polar knowledge scores. To explore this possible difference, the sample was divided into those with pre-program scores lower than or equal to -1.5 (the "lower initial polar knowledge score" group; $n = 22$) and those with scores greater than -1.5 (the "higher initial polar knowledge score" group, $n = 79$). As shown in Figure 11, a repeated measures analysis of variance (ANOVA) with time as the within-subjects factor and initial polar knowledge score as the between-subjects factor indicated that there was a statistically significant difference between the two groups, $F(1, 99) = 23.42$, $p < .001$, with the lower initial knowledge group showing statistically significant gains in polar knowledge, $t(78) = 1.05$, $p < .05$, and the higher initial knowledge group showing no change from pre-program to post-program.

Figure 11. Polar Knowledge Scores by Time and Initial Polar Knowledge Scores



Gains in knowledge were further supported by changes in students' responses to two open-ended questions on the knowledge survey which read "List two ways the polar regions have been affected by a changing climate?" and "Name two tools used by researchers in polar regions. What do scientists measure with these tools?" In total, 39 students responded to the first question on both the pre-program survey and post-program survey. As shown in Figure 12, On the pre-program survey, 28% of these youth responded that they "did not know." On the post-program survey, only 8% of youth responded in this way. In total, 59 students responded to the second question on both the pre-program survey and post-program survey. On the pre-program survey, 29% of youth responded that they "did not know." On the post-program survey, only 5% of youth responded in this way.

Figure 12. Percent of Youth Who Responded "Don't Know" to Polar Knowledge Questions by Time



Importantly, many of the youths' post-program survey responses to these two questions reflect key Polar Literacy Principles, as shown in

Table 3.

Table 3. Youths' Reponses to Open-Ended Polar Knowledge Assessment Items

Polar Literacy Principle	Sample Youth Response
Principle 5A. Arctic sea ice is declining at a rapid rate.	"The ice is starting to melt"
Principle 5A-2. The receding ice cover affects the Arctic food webs and the global ocean circulation; however, the long-term impacts are unclear.	"Some animals can't find enough food and they die"
Principle 5E. The effects of climate change at the Poles is directly connected to changes in sea level around the world.	"The glaciers have been melting and the ocean is rising."
Principle 7. New technologies, sensors, and tools – as well as new applications of existing technologies – are expanding scientists' abilities to study the land, ice, ocean, and atmosphere and living creatures of the polar regions.	"Thermometers measure the weather on a daily basis." "Radar [is used] to scan glaciers." "Drills can carve out layers of ice."

Gains in knowledge are also reflected in youths' responses to open-ended survey items tapping youths' perceptions of what they learned about polar science or the work that polar scientists do. Youth were especially like to describe changes in three areas.

Greater understanding of the characteristics of the polar regions.

> *"I didn't know that there was 24-hour sunlight."*
>
> *"I understand how bad global warning is,"*
>
> *"I learned the different areas of Alaska."*
>
> *"I learned about glaciers."*

Greater understanding of and appreciation for the plight of polar animals.

> *"I learned about the animals."*
>
> *"It changed my thoughts on the animals."*
>
> *"Now I know that Adele penguin population is decreasing."*
>
> *"This program changed my understanding about penguins."*
>
> *"I am knowing now how much danger some penguins are in due to climate change."*

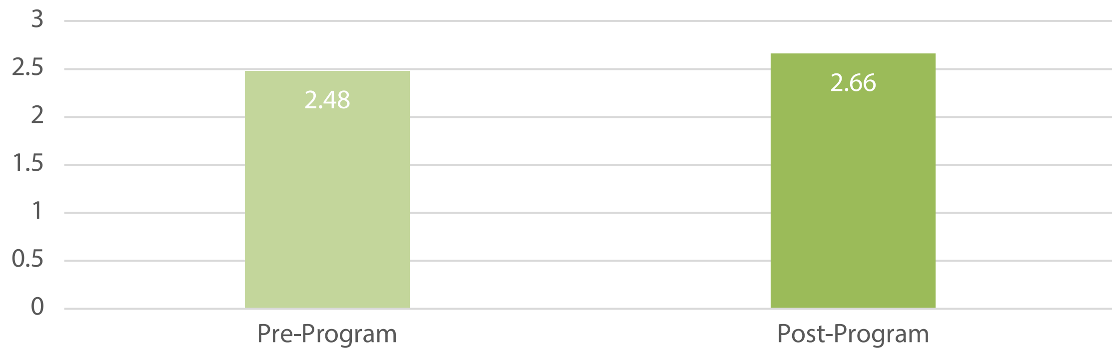
Greater understanding of and appreciation for the work that polar scientists do.

> *"It made me more knowledgeable about what [polar scientists] do."*
>
> *"[Polar scientists] check the temperatures."*
>
> *"I am just now knowing how dangerous their job really is, but also awe inspiring."*
>
> *"It's very important, and you need skills to be a polar scientist."*
>
> *"Their research is important."*
>
> *"This program [caused] me to want to learn more about polar scientists."*

Outcome 5. Participating Youth will Report Increases in Confidence in Working with Data

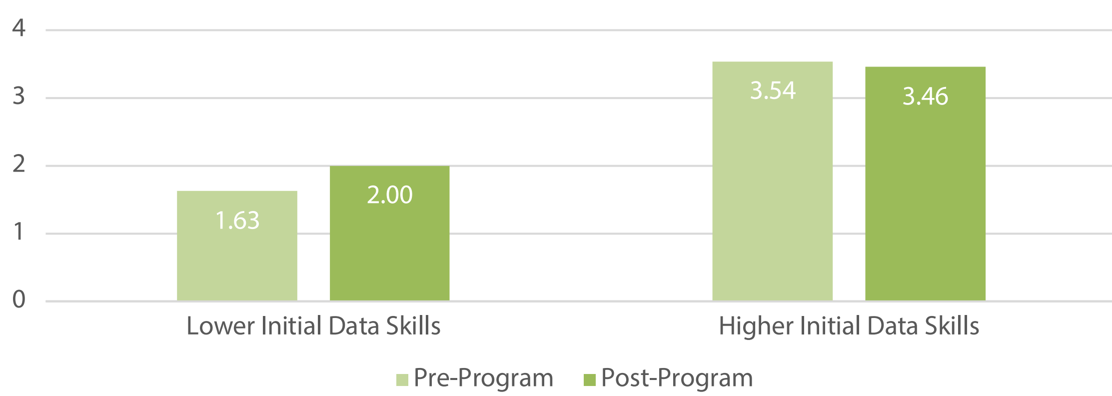
During the Summer 2021 administration, 29 students had both pre-program and post-program data skills scores. As shown in Figure 13, results of a dependent t-test indicated that polar knowledge scores on the post-program survey (mean = 2.66) were higher than those on the pre-program survey (mean = 2.48).

Figure 13. Data Skills Scores by Time



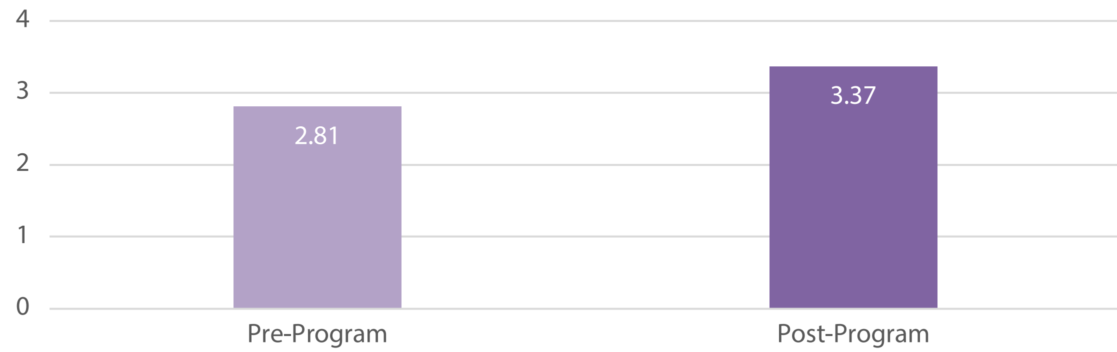
A dependent *t*-test indicated that this difference was not statistically significant. To explore the possibility that gains were stronger among youth with lower initial data skills scores, the sample was divided into those with pre-program scores lower than or equal to 2 (the "lower initial data skills score" group; $n = 16$) and those with scores equal to or greater than 3 (the "higher initial data skills score" group, $n = 13$). As shown in Figure 14, a repeated measures analysis of variance (ANOVA) with time as the within-subjects factor and initial data skills scores as the between-subjects factor indicated that there was a statistically significant difference between the two groups, $F(1, 27) = 4.56$, $p < .001$, with the lower initial data skills group showing statistically significant gains in data skills, $t(15) = 2.09$, $p < .05$, and the higher initial data skills group showing no change from pre-program to post-program.

Figure 14. Data Skills Scores by Time and Initial Data Skills Scores



These findings should be interpreted with caution given small sample sizes. They are, however, consistent with findings from the single item related to confidence in data skills used after Summer 2021. As shown in Figure 15, among the 118 students who completed this assessment, students rated themselves as more confident in "working with data" after the program than before. Results of a dependent t-test indicated that scores post-program ratings (mean = 3.37) were significantly higher than pre-program ratings (mean = 2.81), $t(117) = 5.13$, $p < .001$. The effect size for the increase was $d = .47$ which, according to Kraft (2020), represents a "large" effect compared to other educational interventions.

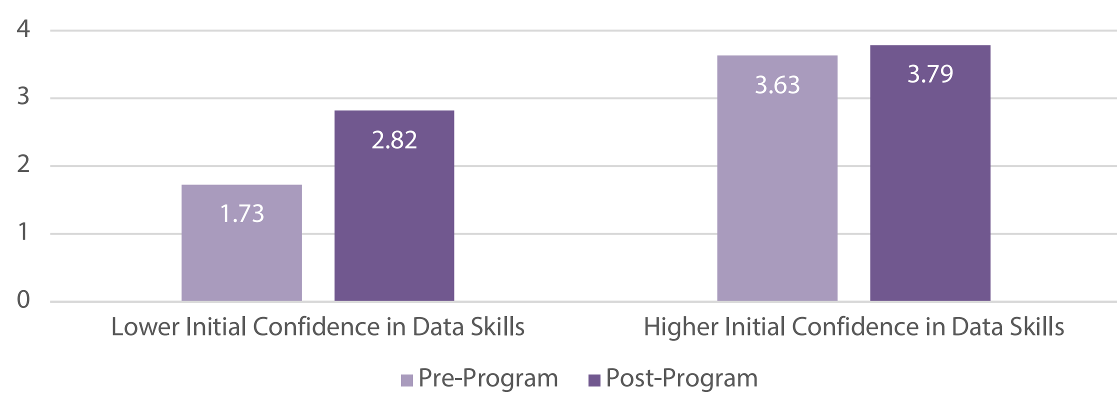
Figure 15. Confidence in Data Skills by Time



Although results of a repeated measured ANOVA indicated that pattern of results did not differ by gender or URM status, it is, again, possible that the pattern of findings differs by youths' initial confidence in data skills. To explore the possibility that gains were stronger among youth with lower initial confidence in their data skills, the sample was divided into those with pre-program ratings lower than or equal to 2 (the "lower initial confidence in data skills" group; n = 51) and those with ratings equal to or greater than 3 (the "higher initial confidence in data skills" group, n = 67). As shown in

Figure 16, a repeated measures analysis of variance (ANOVA) with time as the within-subjects factor and initial confidence as the between-subjects factor indicated that there was a statistically significant difference between the two groups, $F(1, 116) = 20.39$, $p < .001$, with the lower initial confidence in data skills group showing statistically significant gains in confidence, $t(50) = 6.70$, $p < .001$, and the higher initial confidence in data skills group showing no change from pre-program to post-program.

Figure 16. Confidence in Data Skills Scores by Time and Initial Confidence



These findings are important given evidence that, nationally, students lost nearly a full grade level in progress in Data Analysis, Statistics, and Probability scores from 2019 to 2022 as measured by National Assessment of Educational Progress (NAEP). Although these declines can be attributed, in part, to the Covid-19 pandemic, they also reflect longer-term declines in data literacy and data literacy instruction that began well before the pandemic (Drozda, 2022).

The success of the current program in increasing self-reported data skills among students with lower initial data skills as well as confidence in data skills among participants as a whole is consistent with evidence that data skills can be bolstered in youth when instruction includes opportunities for students to create and interact with high-quality visualizations (Krumhansl et al., 2012), uses real-world datasets (Erwin, 2015), and is problem-based and student-directed (English & Watson, 2015). All three strategies were important components of Polar Literacy programming. For example, in the *Data to the Rescue: Penguins Needs Our Help!* adventure, youth are introduced to several different types of data visualizations (e.g., bar graphs, scatterplots, and line graphs representing time series data), formulate and test research questions using real-world datasets including data on changes over time in the population size of Adelie, Gentoo, and Chinstrap penguins collected near the Palmer Research Station in Antarctica, and present a "Data Jam" project that creatively represents the population data (e.g., by visually representing changes in penguin population size from 1995 to 2010 with each penguin cutout representing 100 breeding pairs).

Figure 17. Sample "Data Jam" Project from *Data to the Rescue: Penguins Need Our Help!*

Conclusions

As hypothesized, youth who participated in Polar Literacy programming reported high levels of **engagement** in science learning activities and statistically significant increases in both self-reported **science identity** and **confidence in data skills**. Among students with lower initial scores, participants also reported statistically significant increases in self-reported **fascination with science** and **polar knowledge**.

Importantly, the results reported here are likely to be underestimates of the potential impact of the program for several reasons.

- First, the Polar Literacy project launched in Summer 2020, necessitating an unexpected and rapid transition to virtual programming because of the Covid-19 pandemic. Early on, many youth participants were inexperienced in using Zoom. Throughout the program, many youth participants did not have access to the technology (e.g., a dedicated computer with stable internet) and physical space (e.g., a quiet room) necessary to fully engage.

- Second, student survey data were collected in "real time" as Polar Literacy instructional materials were being developed and revised by two teams of researchers. Later iterations of Polar Literacy programming might be expected to yield stronger outcomes than earlier iterations as project leaders learned from prior experiences and worked to improve both materials and instructional practices. Unfortunately, this possibility is difficult to test empirically given concomitant differences across time in, for example, the characteristics of students, the format of instruction (e.g., virtual vs. in-person), and the level of engagement of out-of-school time staff.

- Third, survey instruments were modified over time to minimize survey fatigue and to better align with developing Polar Literacy program goals and instructional strategies. As a result, some constructs were measured with just a single item (e.g., confidence in data skills) or a very small set of items (e.g., fascination with science). The decision to limit analyses to items that were consistent across time and program sites allowed "apples to apples" comparisons over the course of the project. However, single-item measures can be problematic in several ways (but see Allen, Iliescu, & Greiff, 2022). First, single-item measures are often less reliable (i.e., consistent) than multi-item measures which can result in measurement error. Second, single-item measures may not be as valid (e.g., accurate) as multi-item measures as they may not cover all aspects of the construct being measured. Finally, single-item measures may be less sensitive to change over time than multi-item measures, making it more difficult to detect differences in the construct.

- Fourth, sample sizes were often small. Studies with small sample sizes can be problematic in several ways. First, the responses of small samples of respondents may not accurately represent the larger population from which the sample was drawn, leading to misleading conclusions. Second, small sample sizes can result in a lack of statistical power to detect changes in student knowledge or attitudes over time. Third, small sample sizes make it difficult to examine whether results might differ over time or across groups. In the current study, an attempt was made to examine differences in outcomes by student characteristics (e.g., gender or URM status). However, these results should be interpreted with caution given very small sample sizes.

Recommendations

Based on findings from a review of program materials, survey responses, interviews with educators conducted by evaluators at the Science Education Resource Center (SERC), and observations, the UEPC offers the following recommendations.

1 Efforts should be made to empirically test the effectiveness of specific curricular materials and instructional practices.

Because curricular materials and instructional practices were being developed in "real time" and because sample sizes were small, the current evaluation focused on global changes in students' attitudes and knowledge from pre-program to post-program. Much remains to be learned about which curricular materials and instructional practices are most effective in improving student outcomes. A preliminary analysis of some of these factors is provided in Addendum 2. Quasi-experimental or experimental studies that systematically manipulate curricular materials and instructional practices will be an important next step. Cognitive (or "think-aloud") interviews with youth, informal educators, and/or STEM and informal education professionals might also be utilized to identify areas where, for example, visualizations, wording, or the sequencing or difficulty of activities might be revised to improve clarity and enhance impact.

An initial set of interviews that were conducted by external evaluators at the Science Education Resource Center (SERC) with six educators who implemented Polar Literacy programming provides preliminary evidence that opportunities to interact with polar scientists and to work directly with "real-world" data may be especially impactful. One interviewee noted that the structure of the curriculum was oftentimes quite effective. For example, students were introduced to three different types of penguins before they were asked to grapple with data on changing penguin populations. The perceived result was that the activity was "almost a call to action. 'Now that you love this animal, here is the data about what's happening to it.'" At the same time, these educators indicated areas for improvement. For example, some educators indicated that students lacked the necessary language, math, and science skills (e.g., an understanding of the term 'latitude' or experience with topographical maps) to effectively engage with the curriculum in the timeframe allotted, that asynchronous interactions with polar scientists via pre-recorded videos were less effective than face-to-face interactions, and that CODAP was difficult to use and might need to be replaced with a less sophisticated alternative.

2 Efforts should be made to develop and evaluate the effectiveness of support materials for informal educators.

A recent report released by Data Science 4 Everyone (DS4E; Drozda, 2022) indicates that data literacy scores have declined significantly for both 4th graders and 8th graders over the past decade. In 2022, a national sample of fourth graders had the data literacy skills of third graders from a decade ago while eighth graders had the data literacy skills of sixth graders from a decade ago. Data literacy skills are especially low among low-income students and students who identify as underrepresented minorities (e.g., Black or African American, Hispanic or Latino, or American Indian or Alaska Native). One – of several – likely contributors to these results is a concomitant decline in the emphasis elementary and middle school teachers have placed on instruction related to data literacy.

The Polar Literacy project is already developing a Facilitator Guide for the *Data to the Rescue: Penguins Need Our Help!* adventure to support classroom teachers and informal educators in incorporating data literacy activities into classrooms and out-of-school-time (OST) programs. Surveys, interviews, and cognitive (or "think aloud" interviews) with teachers, informal educators, and STEM and informal education professionals will be helpful in determining whether these materials are meeting educators' needs and what additional resources (e.g., face-to-face professional development) may be necessary. In interviews, several educators indicated that they would benefit from additional supports including instruction on how to utilize CODAP effectively and a "glossary" of some of the specialized language used in the program.

3 Efforts should be made to create curricula and support materials that can be adapted to meet the needs of diverse audiences.

Registration, survey, interview, and observation data all indicated that youth who participated in Polar Literacy programming differed markedly in their backgrounds, initial attitudes, and prior knowledge. Analyses of survey data indicated that youth who began the program with less interest in and knowledge about science were especially likely to report increases in these areas. However, interviews with educators indicated that many of these students also experienced significant challenges. For example, in mixed-age groups, younger children were often unable to complete activities and hearing-impaired students were unable to engage effectively with videos. One interviewee noted the importance of ensuring that educators implementing the programs have the knowledge, time, and support they need to identify and correct student errors, indicating, for example, that students could learn how to "make a graph that appears right, but isn't." This same educator noted the importance of ensuring that curricula are appropriately scaled so that educators and students have time to build rapport and so that educators have time to get a sense of students' personal and educational goals. Professional development sessions may be valuable for helping less experienced educators learn when and how to adapt curricula to meet the needs of the youth with whom they work. An important component of these sessions may be to give educators a chance to learn from one another. For example, in an interview, one educator stressed the importance of introducing youth participants to scientists and to the effects of climate change in their own communities.

References

Allen, M. S., Iliescu, D., & Greiff, S. (2022). Single item measures in psychological science: A call to action [Editorial]. *European Journal of Psychological Assessment, 38*(1), 1–5.

Bakker, A., Cai, J., English, L., Kaiser, G., Mesa, V., & Van Dooren, W. (2019). Beyond small, medium, or large: points of consideration when interpreting effect sizes. *Educational Studies in Mathematics,* 102, 1–8.

Basu, S. and Barton, A. (2007). Developing a sustained interest in science among urban minority youth. *Journal of Research in Science Teaching, 44*, 466-489.

Cole, S. (2012). *The Development of Science Identity: An Evaluation of Youth Development Programs at the Museum of Science and Industry, Chicago.* Dissertations. 339.

Drozda, Z. (2022). *Data science is vital to student success. So why are outcomes going down.* Data Science for Everyone.

English, L.D., & Watson, J.M. (2015). Statistical literacy in the elementary school: Opportunities for problem posing. In: Singer, F., F. Ellerton, N., Cai, J. (eds) *Mathematical Problem Posing. Research in Mathematics Education.* Springer, New York, NY.

Erwin, R. W. (2015). Data Literacy: Real-World Learning Through Problem-Solving With Data Sets. *American Secondary Education*, *43*(2), 18–26. http://www.jstor.org/stable/43694208

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*(4), 241-253.

Krumhansl, R., Peach, C., Foster, J., Busey, A., and Baker, I. (2012). Visualizing Oceans of Data: Educational Interface Design. Waltham, MA: Education Development Center, Inc.

Pfirman, S., Hamilton, L., Turrin, M., Narveson, C., & Lloyd, C. A. (2021). Polar knowledge of US students as indicated by an online Kahoot! quiz game. *Journal of Geoscience Education, 69*(2), 150-165.

Schinske, J., Perkins, H., Snyder, A., and Wyer, M. (2016). Scientist spotlight homework assignments shift students' stereotypes of scientists and enhance science identity in a diverse introductory science class. *CBE - Life Sciences Education 15*(3).

Addendum 1: Analyses of Data from Winter 2023 Implementations

Overview

In Winter 2023, more than 600 additional youth had an opportunity to participate in Polar Literacy programming in four implementations across ten states (Alaska, Arkansas, Florida, Iowa, Kansas, Michigan, Minnesota, New Jersey, Pennsylvania, and Washington). Most of these youth participated as part of their regular, in-person classroom instruction, with just 15 of these youth participating as part of a free choice activity offered by an afterschool program for girls. All participants utilized curricular materials developed as part of the Polar Literacy grant (e.g., Data to the Rescue: Penguins Need Our Help). For all classroom implementations, students had the opportunity to interact virtually with a polar scientist in a 30-minute call.

Sample

All analyses of survey data are based on the 192 youth participants who assented to participate in evaluation activities and who also had parental consent. Among the 172 youth for whom age data were available, the median age was 13 years (range = 9 to 15 years). Among the 154 youth for whom race/ethnicity data were available, 46 (or 29.9%) identified as underrepresented minorities (i.e., Black or African American, Hispanic or Latinx). Among the 162 youth for whom gender data were available, 48.5% identified as female.

Measures

1. Science Identity. Science identity was measured at two time points using four items modified from Cole (2012). Pre-program survey items were administered at the beginning of programming. Post-program survey items were administered at the end of programming. Youth were asked to respond to each item (e.g., "I am interested in pursuing a career in a scientific field") on a five-point scale that ranged from 1 ("not at all true") to 5 ("really true"). Scores on each item were averaged so that higher scores indicated stronger science identity. The four items formed a reliable scale with Cronbach's alphas (αs) ranging from .76 to .80 across the two time points.

2. Fascination with Science. Fascination with science was measured at two time points (i.e., at the beginning and end of the program) using eight items from the Fascination in Science subscale of the Science Learning Activation Survey (2016).[12] Youth were asked to respond to each item (e.g., "I want to know how to do everything that scientists do") on a four-point scale that ranged from 1 (e.g., "NO!") to 4 ("YES!"). Youths' responses were averaged across the items so that higher scores indicated greater fascination in science. The eight items formed a reliable scale with Cronbach's alphas (αs) ranging from .87 to .91 across the two time points.

3. Polar Knowledge. Youth participants were asked to complete a Polar Knowledge assessment at two time points (i.e., at the beginning and end of the program). The assessment includes four items designed by Polar Literacy program staff and modeled on items developed for a national assessment of students' polar knowledge (Pfirman et al., 2021). Each item offers three response options. For the first two items, youth could select more than one response. For these items, each correct response (i.e., selecting a correct response or not selecting an incorrect option) was awarded .5 points while each incorrect selection (i.e., failing to select a correct option or selecting an incorrect option) was assigned -.5 points. For the third and fourth items, youth could select only one response. For these items, a correct response was awarded 1.5 points and an incorrect response was assigned -1.5 points.

[12] See http://activationlab.org/wp-content/uploads/2018/03/Fascination-Report-3.2-20160331.pdf

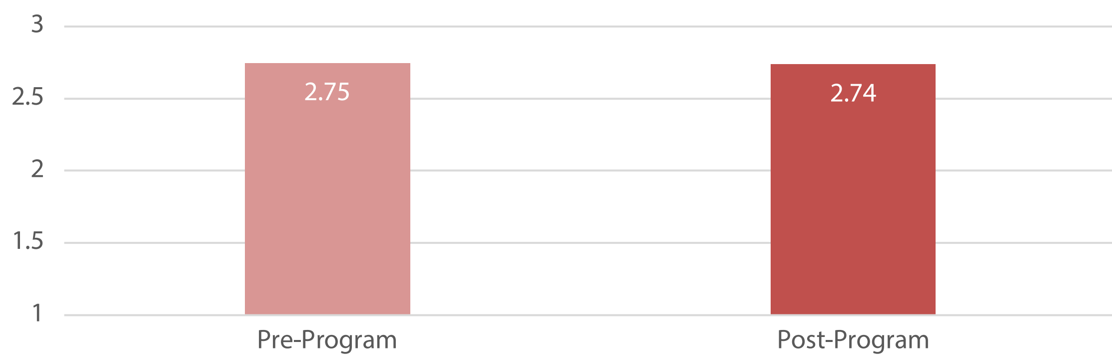
Thus, the possible score for each item ranged from -1.5 to 1.5. Scores were summed across items so that higher scores (ranging from -6 to +6) represented stronger polar knowledge.

Findings

Science Identity

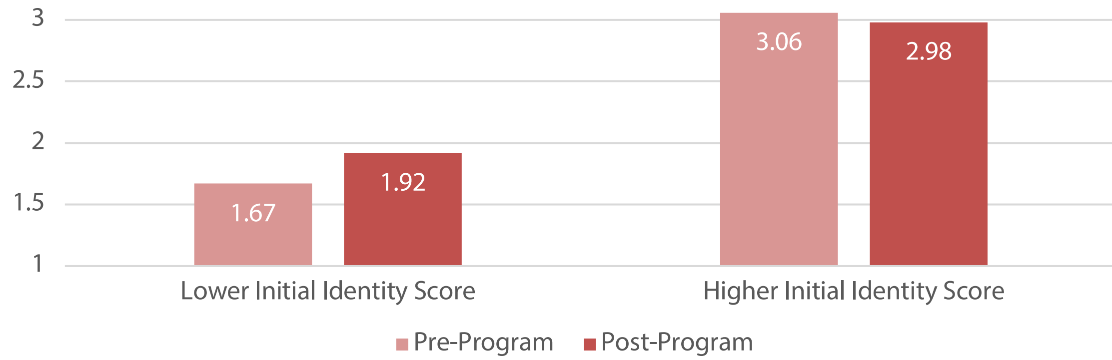
Across implementations, 122 students had both pre-program and post-program science identity scores. As shown in Figure 18, results of a dependent *t*-test indicated that there was no change in youths' science identity scores from pre-program (mean = 2.75) to post-program (mean = 2.74), $t(121) = 0.90$, $p > .05$. A repeated measures analysis of variance (ANOVA) with time as the within-subjects factor and age, gender, and URM status as between-subjects factors yielded no significant effects other than time, meaning that science identity scores and the impact of program participation on science identity scores did not differ by age, gender, or URM status.

Figure 18. Science Identity Scores by Time



To explore whether the impact of Polar Literacy programming might differ for youth with higher versus lower initial science identity scores, the sample was divided into those with pre-program scores lower than or equal to 2 (the "lower initial identity score" group; $n = 27$) and those with scores greater than 2 (the "higher initial identity score" group, $n = 95$). As shown in Figure 19, a repeated measures analysis of variance (ANOVA) with time as the within-subjects factor and initial science identity score as the between-subjects factor indicated that there was a statistically significant difference between the two groups, $F(1, 120) = 6.91$, $p < .01$; only youth in the lower initial identity score group showed gains in science identity over time, $t(26) = 2.57$, $p < .05$.

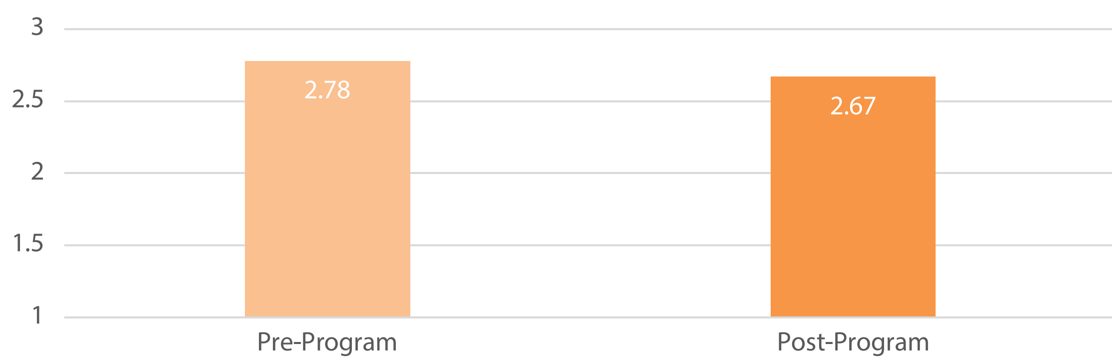
Figure 19. Science Identity Scores by Time and Initial Science Identity Score



Fascination with Science

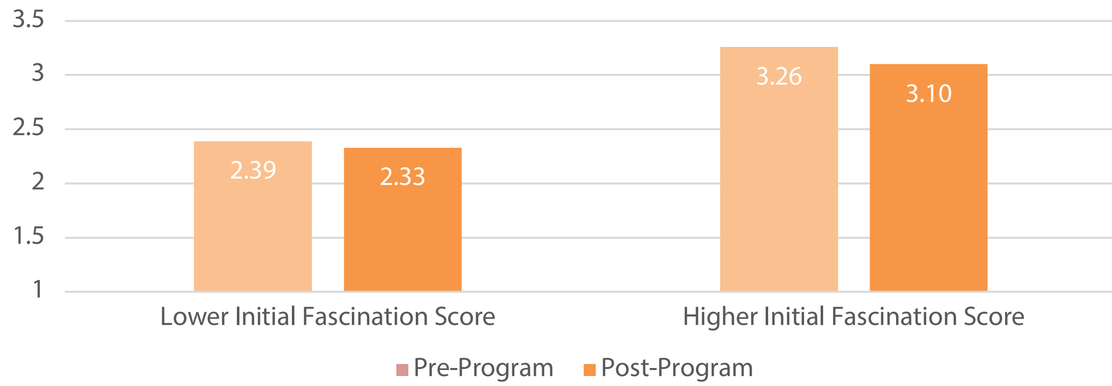
Across sites, 120 students had both pre-program and post-program fascination with science scores. As shown in Figure 20, results of a dependent *t*-test indicated that there was a statistically significant *de*crease in youths' fascination scores from pre-survey (mean = 2.78) to post-survey (mean = 2.67), $t(119) = 2.86$, $p < .01$. A repeated measures analysis of variance (ANOVA) with time as the within-subjects factor and age, gender, and URM status as between-subjects factors yielded no significant effects other than time, meaning that fascination with science scores and the impact of program participation on fascination with science scores did not differ by age, gender, or URM status.

Figure 20. Fascination with Science Scores by Time



To explore whether the impact of Polar Literacy programming might differ for youth with higher versus lower initial fascination with science scores, the sample was divided into those with pre-program scores lower than or equal to 2.75 (the "lower initial identity score" group; $n = 66$) and those with scores greater than 2.75 (the "higher initial identity score" group, $n = 54$). As shown in Figure 21, a repeated measures analysis of variance (ANOVA) with time as the within-subjects factor and initial fascination with science score as the between-subjects factor indicated that there was no statistically significant difference between the two groups, $F(1, 118) = 2.04$, $p > .05$, although follow-up *t*-tests indicated that the decline was significant only among those with initially higher fascination in science scores.

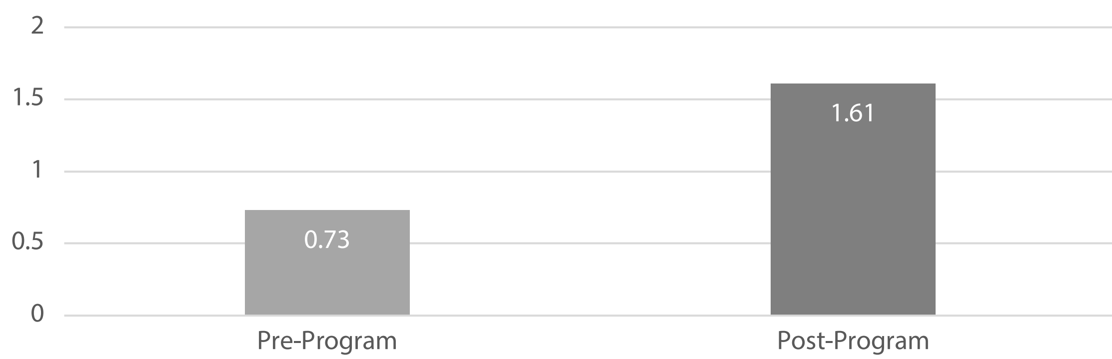
Figure 21. Fascination with Science Scores by Time and Initial Fascination with Science Scores



Polar Knowledge

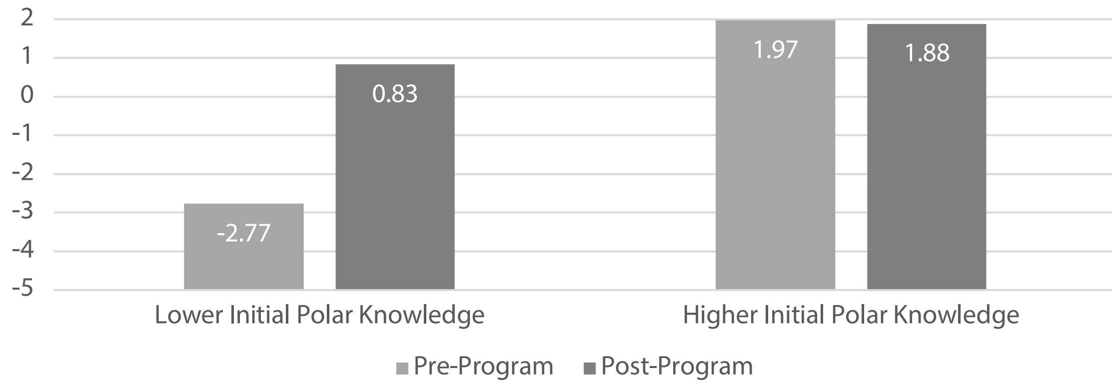
Across sites, 126 youth had both pre-program and post-program Polar Knowledge scores. As shown in Figure 22, the results of a dependent *t*-test indicated that polar knowledge scores at post-program (mean = 1.61) were significantly higher than those at pre-program (mean = .73), $t(125) = 3.08$, $p < .001$. A repeated measures analysis of variance (ANOVA) with time as the within-subjects factor and age, gender, and URM status as between-subjects factors yielded no significant effects other than time, meaning that polar knowledge scores and the impact of program participation on polar knowledge scores did not differ by age, gender, or URM status.

Figure 22. Polar Knowledge Scores by Time



To explore whether the impact of Polar Literacy programming might differ for youth with higher versus lower initial fascination with science scores, the sample was divided into those with pre-program scores lower than or equal to -1.5 (the "lower initial polar knowledge score" group; $n = 33$) and those with scores greater than -1.5 (the "higher initial polar knowledge score" group, $n = 93$). As shown in Figure 23, a repeated measures analysis of variance (ANOVA) with time as the within-subjects factor and initial polar knowledge score as the between-subjects factor indicated that there was a statistically significant difference between the two groups, $F(1, 124) = 42.86$, $p < .001$, with the lower initial knowledge group showing statistically significant gains in polar knowledge, $t(32) = 7.93$, $p < .001$, and the higher initial knowledge group showing no change from pre-program to post-program.

Figure 23. Polar Knowledge Scores by Time and Initial Polar Knowledge Scores



Conclusions

The results of analyses focusing on data collected during Winter 2023 program implementations yielded the following key findings.

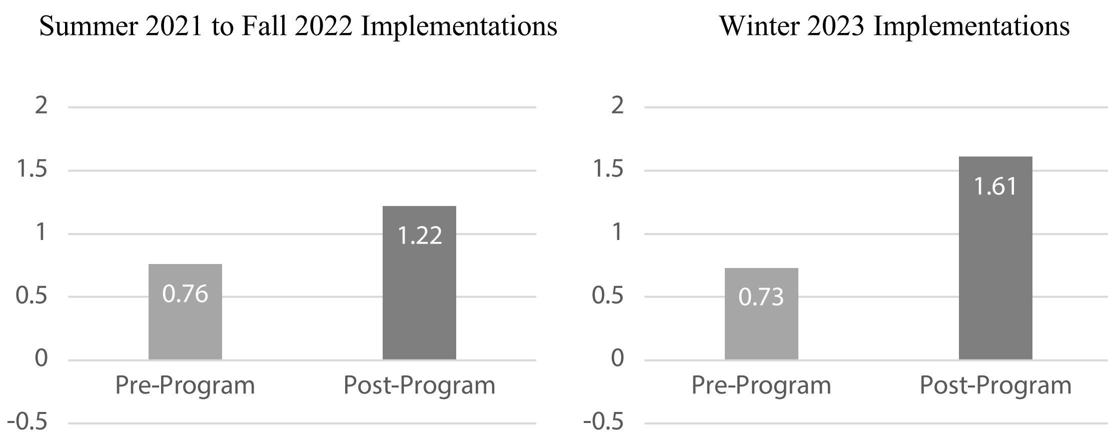
- Youth who participated in Polar Literacy programming in Winter 2023 did not, as a group, report changes from pre-program to post-program in science identity. However, among youth with lower initial science identity scores, there was a statistically significant increase in science identity scores.
- Youth who participated in Polar Literacy programming in Winter 2023 showed small *declines* in fascination with science from pre-program to post-program. These declines were statistically significant only for youth who began the program with initially higher fascination with science scores.
- Youth who participated in Polar Literacy programming in Winter 2023 showed statistically significant increases in polar knowledge. These gains were attributable to gains made by students who began the program with lower initial polar knowledge.

Together, these findings provide some initial insights into how implementation contexts and strategies might affect program impact. During Winter 2023 implementations, most youth engaged in Polar Literacy programming *in person* and as part of their *regular, required classroom instruction.* Moreover, most youth had the opportunity to interact with a polar scientist who participated in a virtual, 30-minute session with youth. In contrast, many youth who engaged with Polar Literacy programming from Summer 2020 to Fall 2022 participated *virtually* and *voluntarily* (e.g., as part of a *free choice* activity associated with an out-of-school-time program). Moreover, some students did not have an opportunity to interact with a polar scientist.

These differences may explain some of the differences in findings between the Summer 2020 to Fall 2022 implementations and the most recent Winter 2023 implementations. Although direct comparisons are difficult for science identity and fascination with science outcomes as the number of items included in measures differed across these two time periods (e.g., fascination with science scores were measured with two to eight items over time), direct comparisons can be made for Polar Knowledge items which remained consistent from Summer 2021 through Winter 2023.

As shown in Figure 24, although students' initial Polar Knowledge was similar across the two time periods, gains were stronger in Winter 2023 compared to prior implementations, perhaps reflecting the benefits of in-person learning or the benefits of being able to interact directly with a polar scientist. As noted above, to fully explore these and other possibilities, quasi-experimental or experimental studies that systematically manipulate curricular materials and instructional practices will be useful. The variation in implementation strategies across sites does, however, allow us to look at some preliminary associations between student outcomes and specific implementation approaches (e.g., whether youth had an opportunity to interact with a polar scientist). We do so in Addendum 2.

Figure 24. Comparison of Change in Polar Knowledge Scores in Summer 2021 to Fall 2022 Implementations vs. Winter 2023 Implementations



Addendum 2: Preliminary Analyses Examining the Impact of Specific Implementation Strategies on Youth Outcomes

Project teams provided additional data regarding implementation that included:

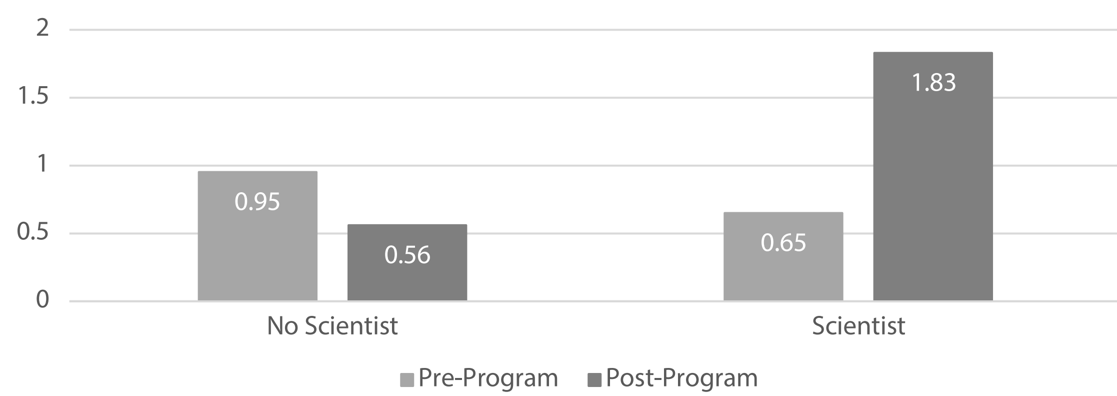
1. the number of contact hours with youth
2. whether the program was offered virtually or in person
3. whether youth had an opportunity to interact with a polar scientist
4. whether the program was required (e.g., part of regular classroom instruction or a required summer program) or voluntary (e.g., part of a voluntary afterschool or summer program).

To explore whether the impact of Polar Literacy programming might differ across these four implementation variables, a series of repeated measures analyses of variance (ANOVAs) and regression analyses were conducted. These analyses focused on just two outcomes – Polar Knowledge and Science Identity – as these constructs were measured consistently enough over time to yield reasonably large sample sizes for these analyses.

The Impact of Implementation Strategies on Increases in Polar knowledge

Analyses indicated that one implementation strategy – whether youth had an opportunity to interact with a polar scientist – had a significant impact on the degree to which students showed increases in scores on the polar knowledge assessment over time, $F(1, 225) = 10.42$, $p < .001$. As shown in Figure 25, among youth who did not have an opportunity to interact with a polar scientist, there was no change in polar knowledge from pre-program to post-program, $t(69) = 0.91$, $p > .05$. Among youth who did have an opportunity to interact with a polar scientist, however, there were statistically significant increases in polar knowledge from pre-program to post-program, $t(156) = 4.49$, $p < .001$. The effect size for these gains was $d = .36$ which, according to Kraft (2020), represents a “large” effect compared to other educational interventions.

Figure 25. Comparison of Change in Polar Knowledge Scores by Presence of a Polar Scientist



The Impact of Implementation Strategies on Increases in Science Identity

Analyses indicated that one implementation strategy – whether the program was required (e.g., part of regular classroom instruction or a required summer program) or voluntary (e.g., part of a voluntary afterschool or summer program) – had a statistically significant impact on the degree to which students showed increases in science identity scores over time, $F(1, 317) = 4.11$, $p < .05$. As shown in Figure 26, among youth who participated in required programming, there was no change in polar knowledge from pre-program to post-program, $t(135) = 0.53$, $p > .05$. Among youth who voluntarily participated, however, there were statistically significant increases in science identity scores from pre-program to post-program, $t(182) = 2.42$, $p < .05$. The effect size for these gains was $d = .18$ which, according to Kraft (2020), represents a "moderate" effect compared to other educational interventions. As might be expected, youth who participated in the program voluntarily had higher science identity scores than youth who were required to participate in programming at both time points, $F(1, 317) = 26.70$, $p < .001$.

Figure 26. Comparison of Change in Science Identity Scores by Degree of Choice