Data Dispositions

One of the goals of the evaluation is to examine the impact of the project on participants' attitudes and dispositions towards data science. During the first year of the project those terms were defined and a tool to measure them was developed. The evaluator collaborated with the project team to define and articulate data dispositions relevant to the Data Clubs project. Based on conversations with the team and a review of data science education literature, a few themes were initially identified. These included: comfort with using a large data set; a sense of curiosity about what story the data could tell; perseverance to keep asking questions; and the desire to reflect upon patterns and observations in the data. These translated into observable behaviors such as technical skills or data moves in CODAP, asking questions iteratively, describing patterns in the data, and communicating findings, among other things.

Once this general outline of the dispositions and skills was created, the team mapped similarities to the work of the Learning Activation Lab[3] (activationlab.org). Activation Lab is a "research and design effort to learn and demonstrate how to activate children in ways that ignite persistent engagement in science, technology, engineering, art, and mathematics learning and innovation." The Lab has created tools for assessing middle school aged students' Science Learning Activation as well as STEM Learning Activation during out-of-school time. Upon closer look, many of the items on these scales aligned with Data Clubs above conceptualization of data dispositions. As the Activation Lab did not have a scale specific to data science, the evaluator and PI worked to adapt the dispositions in this project into a similar tool.

The data dispositions tool was collected before and immediately after youth participation in Data Clubs modules (with the exception of the Covid module, where a modified tool was used). The items are grouped into three categories, similar to the Activation Lab surveys; Fascination, Competencies, and Values.

On all items, there are 26 matched pairs of data, with the exception of a few with only 25, which are noted.

Fascination

---

[3] Chung, J., Cannady, M. A., Schunn, C., Dorph, R., & Bathgate, M., (2016) Measures Technical Brief: Fascination in Science. Retrieved from: http://www.activationlab.org/wp-content/uploads/2016/02/Fascination-Report-3.2-20160331.pdf

Fascination, in this context, could be described as a combination of curiosity, engagement, interest, and persistence. Results from the Fascination items show that most participants find data interesting, like doing projects with data, like making sense of graphs, and are curious about the data behind things that they see in the news. For the majority of these items, the statements were true both before and after Data Clubs. In particular items, however, there were changes from pre to post worth mentioning:

→ On the statement: "After an activity where I work with data, I keep thinking about it", the percentage of those agreeing (yes or YES!) rose from 54% to 77%.

→ For the item: "When I see a headline about a scientific finding, I am curious about the data it is based on", the percentage of those agreeing rose from 69% to 85%.

A. I FIND DATA TO BE:

n=26	Very interesting (1)	Interesting (2)	Boring (3)	Very Boring (4)
Pre	8%	89%	4%	0%
Post	23%	65%	8%	4%

B. IN GENERAL, WHEN I'M DOING A PROJECT THAT INVOLVES DATA, I:

n=26	Love it (1)	Like it (2)	Don't like it (3)	Hate it (4)
Pre	4%	89%	8%	0%
Post	12%	73%	15%	0%

C. AFTER AN ACTIVITY WHERE I WORK WITH DATA, I KEEP THINKING ABOUT IT

n=26	NO! (1)	no (2)	yes (3)	YES! (4)
Pre	0%	46%	46%	8%
Post	4%	19%	65%	12%

D. I LIKE MAKING SENSE OF GRAPHS

n=26	NO! (1)	no (2)	yes (3)	YES! (4)
Pre	4%	12%	58%	27%

Post	0%	8%	65%	27%

E. I'D RATHER HAVE SOMEONE GIVE ME THE CONCLUSION THAN SPEND TIME MAKING SENSE OF A SET OF DATA OR GRAPH MYSELF

n=26	NO! (1)	no (2)	yes (3)	YES! (4)
Pre	15%	54%	23%	7%
Post	15%	50%	35%	0%

F. WHEN I SEE A HEADLINE ABOUT A SCIENTIFIC FINDING, I AM CURIOUS ABOUT THE DATA IT IS BASED ON

n=26	NO! (1)	no (2)	yes (3)	YES! (4)
Pre	4%	27%	46%	23%
Post	0%	15%	62%	23%

Competencies

In this context, Competencies refer to comfort, confidence and knowledge of a variety of skills and approaches to data science. Similar to results on the Fascination items, participants in Data Clubs entered the program with a large degree of competence in their own ability to work with large datasets, interpret graphs, explore problems without one answer, learn new technology, and coming up with questions to explore the data.

Surprisingly, two items, on comfort of working with a large amount of data, and comfort exploring problems without a definite answer, had a smaller percentage of participants agreeing on the post survey than the pre. It is possible that these students came in confident that they would be comfortable with these tasks, however had never actually been exposed to them, and thus realized they are more challenging than first assumed.

A higher percentage of participants agreed that they can understand visual data displays (81% pre to 90% post) after Data Clubs. Additionally, more participants enthusiastically agreed (YES!) that they are comfortable coming up with questions to use to explore data (8% to 23%).

G. I AM COMFORTABLE WORKING WITH A LARGE AMOUNT OF DATA TO TRY TO UNDERSTAND IT

n=26	NO! (1)	no (2)	Yes (3)	YES! (4)
Pre	0%	19%	46%	31%
Post	4%	19%	58%	19%

H. I CAN UNDERSTAND MOST VISUAL DISPLAYS OF DATA (CHARTS & GRAPHS) THAT ARE INTENDED FOR KIDS MY AGE

n=26	NO! (1)	no (2)	Yes (3)	YES! (4)
Pre	0%	15%	54%	31%
Post	0%	8%	54%	39%

I. I AM COMFORTABLE EXPLORING COMPLEX PROBLEMS THAT MAY NOT HAVE A DEFINITE ANSWER

n=26	NO! (1)	no (2)	Yes (3)	YES! (4)
Pre	0%	15%	58%	27%
Post	4%	27%	50%	19%

M. I AM CONFIDENT THAT I CAN USE OR LEARN TO USE TECHNOLOGY TO CREATE VISUAL DISPLAYS FOR DATA

n=26	NO! (1)	no (2)	Yes (3)	YES! (4)
Pre	0%	15%	50%	35%
Post	4%	8%	50%	39%

N. I PREFER TO HAVE STEP BY STEP INSTRUCTIONS WHEN WORKING WITH DATA, RATHER THAN BEING ABLE TO EXPLORE ON MY OWN

n=26	NO! (1)	no (2)	Yes (3)	YES! (4)
Pre	0%	54%	35%	12%
Post	12%	50%	27%	8%

O. I AM COMFORTABLE COMING UP WITH MY OWN QUESTIONS WHEN EXPLORING DATA

n=26	NO! (1)	no (2)	Yes (3)	YES! (4)
Pre	0%	23%	69%	8%
Post	0%	15%	62%	23%

P. WHEN I CAN'T FIND PATTERNS IN DATA, IT MAKES ME WANT TO KEEP LOOKING AND DIG DEEPER

n=26	NO! (1)	no (2)	Yes (3)	YES! (4)
Pre	4%	31%	54%	12%

Post	4%	23%	58%	15%

Values

The items in the Values section assess participants' understanding of how data is relevant and used in the real world, in their communities, and in science.

Results were mixed in this area. A higher percentage of participants agreed that they understand how data is related to community issues after their participation in Data clubs (89%), than prior to (77%). The percentage enthusiastically agreeing (YES!) that data is necessary for scientific discoveries rose from 54% to 73%, and that working with data is important from 39% to 50%. The final three items did not show the same level of impact.

Q. KNOWING HOW DATA IS USED CAN HELP ME IN UNDERSTANDING ISSUES IN MY COMMUNITY (FOR EXAMPLE, ELECTION RESULTS, WATER QUALITY, SCHOOL RANKINGS)

n=26	NO! (1)	no (2)	Yes (3)	YES! (4)
Pre	0%	23%	58%	19%
Post	0%	12%	62%	27%

R. PEOPLE WHO WORK WITH DATA DO IMPORTANT WORK

n=26	NO! (1)	no (2)	Yes (3)	YES! (4)
Pre	0%	4%	58%	39%
Post	4%	4%	42%	50%

S. DATA IS NECESSARY FOR SCIENTISTS TO MAKE NEW DISCOVERIES

n=26	NO! (1)	no (2)	Yes (3)	YES! (4)
Pre	0%	8%	39%	54%
Post	4%	4%	19%	73%

T. I'D RATHER JUST GOOGLE SOMETHING THAN EXPLORE THE DATA OR GRAPHS ON MY OWN

n=26	NO! (1)	no (2)	Yes (3)	YES! (4)
Pre	4%	54%	31%	12%
Post	19%	58%	23%	0%

J. KNOWING HOW DATA IS USED/PRESENTED WILL HELP ME UNDERSTAND HOW THE WORLD WORKS

n=26	All the time (1)	Most of the time (2)	Sometimes (3)	Never (4)
Pre	19%	43%	33%	5%
Post	10%	38%	48%	5%

K. UNDERSTANDING HOW TO WORK WITH DATA WILL HELP ME DO WELL IN:

n=26	All classes (1)	Most classes (2)	Some classes (3)	None of my classes (4)
Pre	15%	39%	46%	0%
Post	8%	31%	50%	12%

Project Research

A prominent piece of this project was for the project team to conduct research on the implementation and impact of the modules. Throughout the past two years, an interview-based assessment has been developed, piloted, and revised. The interactive interview involves giving the participant a novel dataset and asking them to perform various tasks, determine if particular questions could be answered with the available data, and come up with their own questions to explore. Eleven (11) 30-45 minute interviews were conducted and recorded and are currently being analyzed by the team.

Additionally, the project has produced and disseminated a number of products including a new website (https://www.terc.edu/dataclubs/) where the modules are available for download, a video in the STEM for All Video Showcase, and a number of conference presentations and articles.