

Understanding and Improving Data Visualization Literacy: Compiling Front-end Evaluation and Partial Summative Evaluation Data

Prepared by:
Joe E. Heimlich, Ph.D.
E. Elaine T. Horr, Ph.D.

October 2020

*This project was completed
with support from
the National Science Foundation
(1713567)*



Introduction

This Innovations in Development AISL project is a collaboration of effort between Indiana University (IU), the Science Museum of Minnesota (SMM), the University of California, Irvine, and COSI's Center for Research and Evaluation (formerly known as Lifelong Learning Group, or LLG, and referred to now as CRE). This project is motivated by the insight that in the information age, being able to “read and write” data visualizations, or data visualization literacy, is becoming as important as being able to read and write text.

Sense-making with data through the process of visualization—recognizing and constructing meaning with these data—has been of interest to learning researchers for many years. Results of a variety of data visualization projects in museums and science centers suggest that visitors have a rudimentary understanding of and ability to interpret the data that appear in even simple data visualizations. This project supports the need for data visualization experiences to be appealing, accommodate short and long-term exploration, and address a range of visitors' prior knowledge.

The Run exhibit was designed to be installed at SMM, IU, UCI, COSI, and other possible sites that might have been identified as the project continues. This exhibit asked participants to input some personal data into a computer at the beginning of their experience. They then ran or walked, depending on the site, along a track and sensors recorded starting and ending times. These data as well as the data entered by the participant helped populate a data table and data visualizations that appeared on a monitor at the end of the running/walking track. Participants had the opportunity to manipulate different types of data visualizations using the data they entered at the beginning of their experience as well as data entered from previous visitor participants.

COSI's Center for Research and Evaluation

Front-end evaluation occurred during the spring and early summer of 2019 at IU and at COSI. On the basis of results of this evaluation, SMM and IU made software and hardware changes to the exhibit components and shipped them to COSI in January 2020. After doing some testing of the final exhibit prototype, summative evaluation commenced in late February 2020 and continued through early March 2020 until the pandemic shut-down.

Evaluation Questions & Methods

Front-end Evaluation

During year 2 of the project, CRE developed the overarching question driving the front-end evaluation research is “Is there a difference in engagement from a laboratory setting to in-situ on a museum floor? If so, in what ways?” and the supporting evaluation research questions.

- *How do participants engage with the Run (Walk) exhibit?*
- *What are participants’ immediate reactions to engaging with the Run (Walk) exhibit, especially the data visualization aspect?*
- *What impressions regarding data visualization do participants take with them after engaging with the Run (Walk) exhibit?*
- *To what extent did engaging with the Run (Walk) exhibit encourage participants to “go deeper” into data visualization exploration?*

CRE tested the Walk exhibit in three separate areas during public open hours: for the lab testing data collection was performed in a controlled environment free of noise and distractions; the first in situ testing was done in a fairly isolated and quiet hallway with little foot traffic; and the second in situ testing was done in an area that had high foot traffic and was noisy. We used the same combined observation and interview instrument for all three areas, with the exception of the addition of the think aloud interviews we did with visitors in the lab testing environment. For the think aloud data collection, we wanted visitors to engage with the exhibit and speak aloud what they were thinking as they went through each step of their exhibit experience to better understand why they are engaging in specific actions, especially with the data visualizations.

Summative Evaluation

During year 3 of the project, CRE developed data collection instruments designed to measure affective and cognitive outcomes from the experience along with a measure of transference. These instruments were based on data gathered during front-end studies: the lab testing and museum in situ testing of the initial prototype done during the second project year. Additionally, we inserted items specifically designed to measure the influence of new data visualizations added to the MAV software for the summative testing phase to the instruments. The plan called for collecting consistent outcome data at four sites, COSI, SMM, IU, and UC Irvine, using the final prototype of the MAV exhibit. The overarching question guiding the summative evaluation was “What information is necessary for the ultimate visitor meaning-making of data visualizations?” Supporting evaluation questions include:

- *What personal attributes do visitors place more importance on to be connected to an experience such as the Run exhibit?*
- *To what degree did having the visitor’s personal information as part of the data set that populated the data visualizations offered in the MAV influence their interest in making meaning of the data?*

- What other types of experiences in the museum do the visitors think this type of visualization might enhance?
- To what extent is this understanding transferrable to other data visualizations?

The data collection plan also called for collecting data under four levels of interactivity between the data collector and the visitor. These levels ran the spectrum from the evaluator not recruiting or giving the visitor information about the exhibit prior to the visitor's engagement with the exhibit to the evaluator recruiting the visitor and giving detailed information about the study and the exhibit. The four levels of the evaluator interactivity with the visitor are as follows:

1. The exhibit is on the floor and visitors engage with it without being recruited to do so; interview visitor at end
2. We invite visitors to engage with the exhibit; interview visitor at end
3. We invite visitors to engage with the exhibit and give basic guidance on what to do with the exhibit; interview visitor at end
4. We invite visitors to engage with the exhibit, explain the purpose of the study, and give basic guidance on what to do with the exhibit; interview visitor at end

Results

Front-end Evaluation

We interviewed 156 adult visitors who engaged with the exhibit either in the lab setting, one of the in situ settings, or in the think aloud process. A summary of the overall engagement with this exhibit can be seen below in Table 1. Overall key findings that led to discussions and changes in the MAV included:

- Data input kiosk:
 - Height: some visitors were either not sure of child's height or had trouble converting height in feet and inches into just inches
 - Some steps of the data entry were not intuitive, even with sign explaining what to do, and too much lag time between steps
- The competitive aspect was most important part of the experience.
- MAV engagement:
 - Overall, most frequent engagement was only with the data table.
 - As the experience environment progressed from no distractions (lab testing environment) to on the floor with all the regular distractions, we saw about the same level of interaction with the data table, but less with the graph and geomap.
 - The touch and drag motion for the scatter graph was not at all intuitive, even with sign explaining what to do
 - Most visitors who engaged with the MAV, understood the data that appeared in the data table, scatter graph, and geomap due to prior experience with those types of data visualizations.

Table 1: Summarizing engagement with the exhibit

Testing sessions	# of Children-teens	# of Adults	Shortest MAV time	Longest MAV time	Average MAV time	Looked at data table	Worked with scatter graph	Worked with geomap
			Time in seconds					
Lab	16	40	27	465	176	All	All	All
Think aloud	9	7	n/a	n/a	n/a	All	All	All
1 st in situ	30	49	15	435	132	All	Usually	Sometimes
2 nd in situ	85	60	10	180	48	Usually	Rarely	Rarely

The specific findings by evaluation question follow.

1. How do participants engage with the Run (Walk) exhibit?

- For all settings, almost all the visitors chose to do the walking portion of the experience, with only occasionally adults merely helping to facilitate the data input for younger people in their group and not walking.
- The majority of times, visitors engaging with the exhibit when it was installed in the second in situ setting mainly looked at their time on the MAV data table and did not engage with the scatter graph or geomap.
- Visitors took considerably more time in engaging with the MAV data visualizations when the exhibit was installed in the lab setting and the first in situ setting than when it was installed in the second in situ setting.

2. What are participants' immediate reactions to engaging with the Run (Walk) exhibit, especially the data visualization aspect?

- Visitors who engaged with the Walk exhibit rated their experience slightly higher, based on a 7-point scale, when the exhibit was installed in the lab setting ($\bar{x} = 4.9$) and the first in situ setting ($\bar{x} = 5.0$) than in the second in situ setting ($\bar{x} = 4.7$).
- Participants put high importance on the walking portion of the experience, mainly due to the friendly competitive factor within the group with which they were visiting.
- Visitors showed more interest in engaging with MAV data visualizations when the exhibit was installed in the first two more isolated and quieter areas of the building than in the louder and busier area of the building.
- Visitors told us it was important to have their data appear in the data visualizations – most otherwise would not have had any interest in the MAV data visualizations.

3. What impressions regarding data visualization do participants take with them after engaging with the Run (Walk) exhibit?

- In all the data collection settings, especially the lab and first in situ setting, there were a few visitors who told us they really enjoyed manipulating the MAV data visualizations
- A few visitors noticed that although the geomap had a global version, there was no way to input non-American codes that would indicate where someone was from so if there were visitors from other countries, their geographical data would not appear.
- A very small minority of visitors criticized using a scatter graph rather than other types of visualizations they felt would be more appropriate for the type of data that was being collected.
- Many visitors shared with us that there really was not much to manipulate in the data visualizations and therefore they did not spend much time at the MAV.

4. To what extent did engaging with the Run (Walk) exhibit encourage participants to “go deeper” into data visualization exploration?

- There were very few times visitors engaged with the MAV data visualizations other than in a cursory manner. There were, however, some exceptions.
- Lab setting: one of the young adults attending the after-hours COSI After Dark event spent the longest time of engagement with the MAV data visualization and was identifying and explaining different variable connections on the graph and geomap to people in the group with whom he had attended the event.
- First in situ setting: the group that spent the longest time engaging with the MAV data visualizations, a father and his two daughters, all did the walk and the 13-year-old daughter was showing and explaining the different data visualizations to her 11-year-old sister.
- Second in situ setting: visitors did very little with the MAV data visualizations beyond finding their time on the data table and manipulating the axes in the scatter graphs – very little discussion occurred.

In answer to the overall evaluation question “Is there a difference in engagement for the lab testing and the in-situ testing?”, we would assert that there is a difference. Visitors interacted with the MAV data visualizations more often and for longer time periods in the data collection settings that were more isolated and quieter than in the data collection setting in the hallway that had much higher foot traffic and was much noisier, which is the more typical science center setting for exhibits.

Summative Evaluation: Phase 1

During January and February 2020, we tested the components of the functional final prototype of the MAV prior to starting use of the new data collection instruments for the initial phase of the summative evaluation. After identifying and having small functional problems corrected, we started collecting data in late February 2020 and continued until very early March 2020, after which COSI closed due to the SARS-CoV-2 pandemic. This closure continued through late spring, summer, and is ongoing at the time of this report meaning the summative and the research components of the project are unable to be completed. Likewise, SMM closed and was unable to do their data collection. Therefore, we have been unable to complete any but the most basic summative evaluation data collection for the first phase of this evaluation. While we are not comfortable making transferrable or generalizable claims, the data do present ideas of how some COSI guests engaged with the MAV. Table 2 captures the key observations in the preliminary phase of data collection. One component of the summative study was to have been a comparison of time spent on the MAV with a comparable experience in each of the two museums, thus giving us an honest comparison of the extent to which the MAV increases engagement time for some visitors and the relative proportion of visitors who engage with an activity and a secondary component across visitors.

Table 2: Partial Phase 1 Summative evaluation results (7-point Likert-type scale)

Person by age group; level of recruitment	Mean time engaging with exhibit (minutes)	I like to do this type of thing	People with me like to do this type of thing	Important to see myself in the data	Important to be able to compare with others

Youth, child 8-17 yrs						
Interview only (n=8)	5.04	5.8	3.2	5.9	5.8	
Invite, interview (n=3)	4.00	3.8	3.3	3.8	3.2	
Adult, 18-29 years						
Interview only (n=3)	3.46	6.0	6.0	6.8	6.5	
Invite, interview (n=1)	2.23	4.5	3.5	2.5	3.0	
Adult, 30-49 yrs						
Interview only (n=12)	4.98	5.7	5.9	5.7	5.6	
Invite, interview (n=2)	3.33	3.3	3.8	4.8	2.3	
Adult 50+ yrs						
Interview only (n=1)	-	7.0	7.0	6.5	7.0	
Invite, interview (n=1)	-	5.0	5.0	4.0	1.5	

Main points:

- Youth/children and adults 30-49 years of age spent the most time engaging with the exhibit. This could be explained because that adult age group is likely to have children 8-17 years of age and therefore they would be engaging with the exhibit together as a group.
- Data suggest adults chose to engage with the MAV because they enjoyed that type of activity, slightly more than youth/children did.
- Youth and children appear not to take into consideration whether someone else in their group would be interested in the activity when they chose to engage with it. Although all adult groups did take interest in the activity of others in their group into consideration, data suggest the older the adult is, the more they would take others' preferences into consideration.
- Participants in all age groups indicated it was important to see themselves in the data appearing on the MAV
- Young adults (< 30 years) told us it was extremely important to be able to compare things using the MAV while the oldest adult group (50+ years) told us it was moderately important. The middle age group of adults and the youth/children thought it was important to be able to compare data.

Comparing the results

We intentionally designed the data collection instruments for the front-end and summative evaluations to obtain mostly differing types of data. However, we are able to compare the degree to which visitors participating in this study engaged with the major components of the MAV across both studies, strengthening our trust in the trends revealed (Table 3). Data suggest almost 100% of participants engaged with the data table. Because we did not specifically have an item on the front-end evaluation instrument in which to record data as to how many participants sorted data in the data table, we do not know how many chose to do so. When looking at engagement levels with the scatter graph and the geomap, we found that as the front-end environment become louder and had more distractions, the level of engagement with these two data visualizations decreased. For the summative data we were able to collect, we really were unable to note any patterns of significance for interaction beyond merely looking at their data initially in the data table.

Table 3: Engagement level: front-end evaluation data collection environment; summative evaluation

	Data table		Sorted data		Scatter graph		Geomap	
	Did	Did not	Did	Did not	Did	Did not	Did	Did not
Front-end Evaluation: Lab setting data								
Youth/Child 8-17 years (n=16)	16	0	n/a	n/a	16	0	16	n/a
Adults 18 yrs+ (n=40)	40	0	n/a	n/a	40	0	40	n/a
Front-end Evaluation: In situ data – 1st round (quiet area on the floor)								
Youth/Child 8-17 years (n=30)	30	0	n/a	n/a	Usually	Unknown	Some-times	Unknown
Adults 18 yrs+ (n=49)	49	0	n/a	n/a	Usually	Unknown	Some-times	Unknown
Front-end Evaluation: In situ data – 2nd round (on floor with normal distractions)								
Youth/Child 8-17 years (n=85)	85	0	n/a	n/a	Rarely	Unknown	Rarely	Unknown
Adults 18 yrs+ (n=60)	60	0	n/a	n/a	Rarely	Unknown	Rarely	Unknown
Summative Evaluation: new								
Youth/Child 8-17 years (n=12)	11	1	10	3	6	5	2	10
Adults 18 yrs+ (n=4)	4	0	2	2	3	1	0	4

From these evaluations, it is possible to say that the engagement of the activity works, and the capture of comparative data drives interest. The evaluations suggest that there are challenges with the interaction between the experience (walk/run) and the MAV interface, but that when people do engage, the MAV does lead to reflection, asking questions, and using the graphics to make meaning.