Engaging Museum Visitors with Scientific Data through Visualization: A Comparison of Three Strategies

Joyce Ma
Exploratorium

Annual Meeting of the American Educational Research Association
San Francisco, April, 2013

This study compares three strategies: (1) open-ended exploration, (2) narrative introduction, and (3) challenges; for engaging museum visitors with scientific data through visualization. It considers how the strategies support visitors in investigating the distributions of marine microbes with Living Liquid, a visualization tool designed to allow visitors to look at changing plankton patterns and correlate those patterns with time-varying environmental factors. Analysis of think-aloud transcripts for 94 dyads, randomly assigned to a strategy, indicate that challenges decreased the number of questions generated, and the narrative introduction helped visitors connect the exhibit to the importance of plankton and its research. Findings suggest guidelines for designing visualizations to support exploration of scientific datasets in the informal learning context.

KEYWORDS: visualization, STEM, informal learning, museums, science centers, ocean microbiology

Introduction

Visualization tools in museums can give the public access to the data scientists use, and, like microscopes or telescopes, allow visitors to explore new worlds (Frankel & Reid, 2008; Johnson et al., 2006). However, to be useful learning tools for museum-goers, visualizations need to engage visitors, both domain novices and experts, in investigating data in ways that are personally meaningful and authentic to science, all within the few minutes people spend at a museum exhibit. This study compares three strategies for supporting museum visitors in exploring a scientific dataset using a visualization tool at a museum. The tool, *Living Liquid*, provides museum-goers access to research data from the MIT Darwin Project of the global distributions of marine microbial populations as they change over time.

The three strategies considered are (1) *open-ended exploration* - visitors use *Living Liquid* however they like, (2) *narrative introduction* - visitors watch a short slideshow describing where the dataset came from and its scientific significance; and (3) *challenges* - visitors are asked to investigate questions that can be answered with the dataset in *Living Liquid*. Comparison among the three seeks to answer the question: How do different strategies engage visitors in data exploration? This study looks, in particular, at three aspects of data exploration: (1) thoroughness of use, (2) seeing patterns, and (3) asking and answering questions. The study also compares the interest ratings visitors gave to their experiences at the exhibit.

Background

Visualizations take advantage of the capacity of the human visual system to allow people to engage with information much more quickly than with raw numbers or text (Carroll, 1993; Johnson et al., 2006). As the amount of data scientists collect grows exponentially, visualization is becoming a central tool in many scientific disciplines that allows researchers to interact with the data they collect to make observations, detect patterns, and otherwise enable discovery of our world (Frankel & Reid, 2008; Johnson et al., 2006; Thomas & Cook, 2005). The ability to create and interpret visualizations is increasingly recognized as a critical skill to advancing scientific discovery and understanding (Johnson et al., 2006; Thomas & Cook, 2005).

Informal science organizations are just beginning to explore ways of visualizing scientific data at museums. Some examples include NOAA's Science on a Sphere network, the Future Earth project at the Science Museum of Minnesota, Adler Planetarium's Visualization Laboratory, the Tahoe Environmental Research Center Visualization Lab, and the NISE Visualization Lab. Much of the recent efforts have focused on using visualizations to communicate science content. *Living Liquid*, on the other hand, aims to engage visitors with data by providing a visualization tool for exploration. It is, therefore, more similar to the exhibit, *Rain Table* developed at the Science Museum of Minnesota, that allows visitors to interact with scientific data by selecting rain locations and seeing how rainwater moves across the landscape based on mathematical models used by researchers.

Edelson and Gordin (1998) outlined a framework for adapting visualization tools for learners that provided useful guidelines for the development of *Living Liquid*. However, their framework was formulated largely from classroom studies, while the museum environment, where learning is characterized by unmediated, episodic and short interactions with exhibits (National Research Council, 2009), imposes its own constraints on learning with a visualization tool. Successful strategies for engaging visitors with exploring data through visualization tools in museum settings still need to be identified.

This work compares three candidate strategies for engaging visitors with data exploration using *Living Liquid*. The first strategy, *open-ended exploration*, allows visitors to use the visualization tool with no additional guidance beyond the tool's affordances. Prior work in developing open-ended exhibit experiences have found that visitors are more likely to ask and answer their own questions at these exhibits (Gutwill, 2005; Borun et al. 1998). On the other hand, visitors may not have the background content information to make sense of their explorations or the inquiry skills to pursue productive investigations at such exhibits (Allen & Gutwill, 2009). Since exhibit experiences are rarely mediated by staff, meaningful exploration depends on the design of the visualization tool to support visitors in interpreting the visualization and the data it represents, to spark curiosity in the data, and to guide visitors in answering self-generated questions with the data.

The second strategy, *narrative introduction*, includes a short slideshow describing where the data come from and how the data are used in research. This slideshow is shown to study participants before they use *Living Liquid* and serves to situate the tool within a larger narrative of current scientific research. By telling the story behind the data, it can provide the broader context for the data: what the dataset is (data used by scientists), why it is important to researchers, and how it is used by scientists. The introduction is intended to help visitors make personal connections to the data and reflect on the relevance of the scientific dataset they are about to explore, which in turn may motivate more meaningful exploration. At the same time, prior attempts at introducing narratives to hands-on exhibits at the Exploratorium have found that narratives can be difficult to incorporate into such exhibits, and video stories designed to help visitors find personal significance at exhibits do not readily encourage more exploratory behavior (Allen, 2004).

The third strategy involves asking visitors to try to meet *challenges* (e.g., "*Prochlorococcus* are a type of phytoplankton that do not need much nutrients. Where might you find *Prochlorococcus*?") that can be answered with the dataset in *Living Liquid*. Visitors are also free to pursue their own questions throughout their interactions with the visualization tool. Staff-authored challenges model question-asking by giving examples of the types of questions that are productive to ask of the dataset. Also, by beginning with an easier question, the sequence of challenges may give visitors, unfamiliar with the visualization tool, a chance to learn its basic affordances and the data before looking for more sophisticated relationships (e.g., correlations between multiple variables). Alternatively, asking visitors to pursue staff-authored challenges may detract visitors from asking their own questions while they pursue challenges that might have little meaning to them, a concern in the free-choice learning environment of a museum.

Material and Method

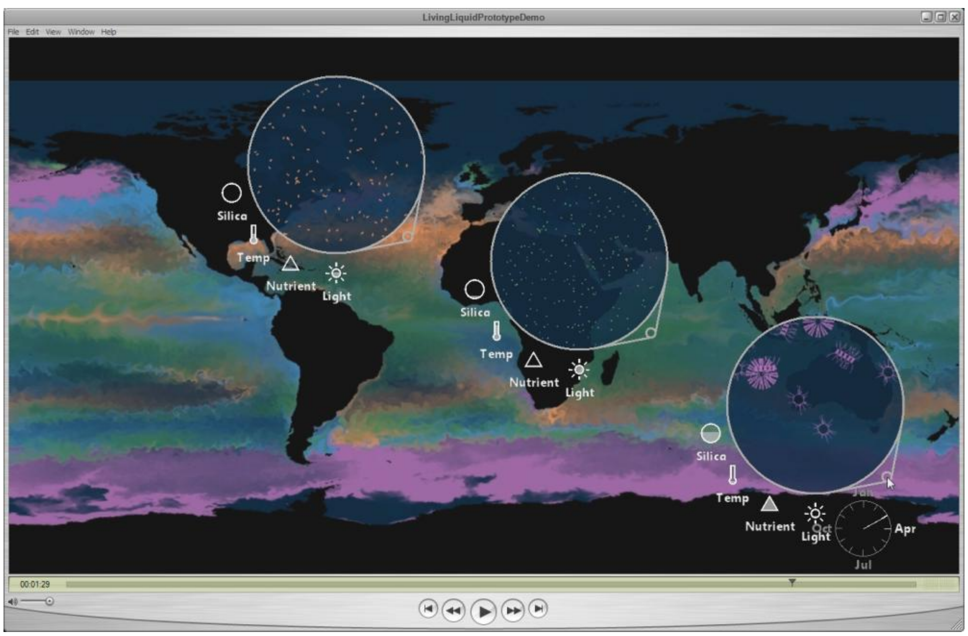


Figure 1. *Screenshot of Living Liquid visualization.*

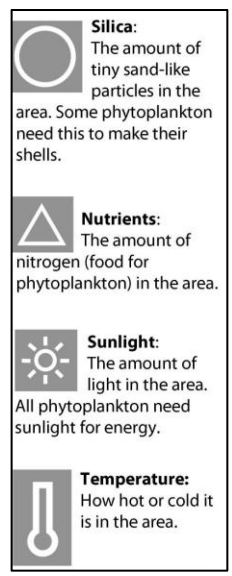


Figure 3. *Legend for the four environmental factors.*

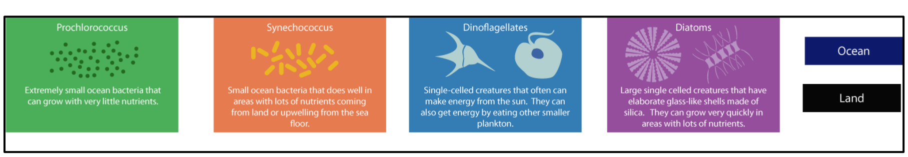


Figure 2. *Label describing the four different plankton types.*

Living Liquid was iteratively developed guided by user-centered design principles by the Exploratorium and the Visualization and Interface Design Innovation (ViDi) group at UC Davis. The version of *Living Liquid* considered in this study used a single-touch 26-inch, touchscreen monitor, large enough to comfortably accommodate two users at a time. It presented visitors with a global map of the world's oceans. Different types of phytoplankton were represented by different hues and each type's relative abundance was encoded in the color's alpha value, which roughly translates to transparency. The changing population distribution of the different phytoplankton types over six years was visualized in a looping animation. Visitors could touch a portion of the map, stopping the animation, and a circle viewer would appear, showing the types of phytoplankton in the area touched. Indicators flanking the circle viewer showed the relative levels of four environmental factors: temperature, sunlight, nutrients, and silica; at that location. Three circle viewers could be brought up at the same time to aid comparison among different locations. Figure 1 shows a screenshot of *Living Liquid*. Static labels were mounted on the periphery of the monitor describing the four plankton types (Figure 2) and the four environmental factors (Figure 3).

Figure 4. *The Living Liquid prototype in the Exploratorium's Life Sciences area.*

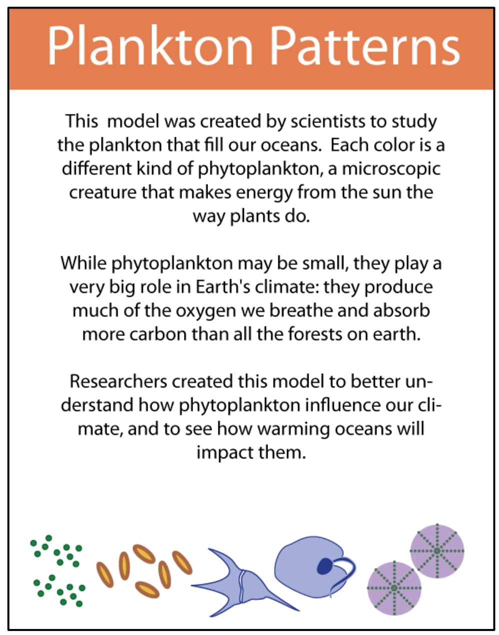


Figure 5. *Exhibit label.*

The prototype was placed in the life sciences area of the museum (Figure 4), along with a sign describing the exhibit's data and its importance (Figure 5). A data collector approached every third visiting pair who crossed a pre-drawn imaginary line and asked for the two visitors' participation in this study. To be asked, both visitors had to be 11-years old or older in keeping with the target audience for the exhibit; if both visitors were minors, the researcher also secured verbal consent from an accompanying adult. The data collector cycled from one strategy to the next, and each dyad experienced only one strategy: *open-ended exploration*, *narrative introduction*, or *challenges*.

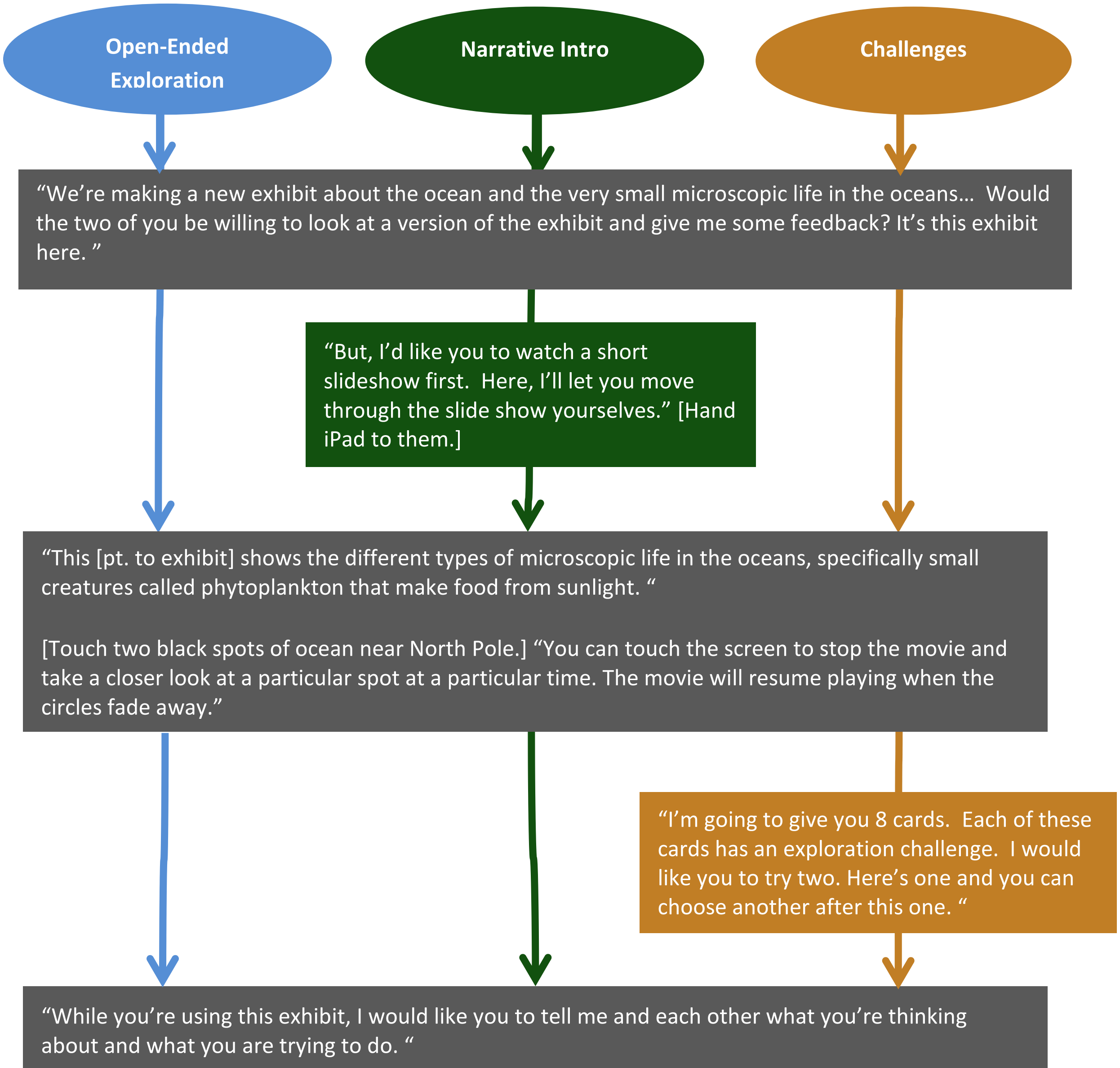


Figure 6. *Instructions given to each dyad for the three strategies.*

The instructions given to a dyad for each of the three strategies are outlined in Figure 6. In *open-ended exploration*, visitors were asked to use the *Living Liquid* prototype however they liked. Alternatively, visitors assigned to the *narrative introduction* strategy, were given a slideshow to look through before they used *Living Liquid*. This slideshow, included in Appendix A, consisted of nine slides that described how and why scientists study plankton and the data that visitors could explore with *Living Liquid*. It was presented on an iPad, and visitors could move through the show at whatever speed they preferred. Afterwards, visitors in the *narrative introduction* group could use *Living Liquid* however they liked.

With the *challenges* strategy, visitors were presented with a set of eight challenge cards, each of which was designed to be answerable with the data available in *Living Liquid*. The cards varied in the type of data they asked visitors to consider: plankton type, seasonal variation, abundance, diversity, and community composition. Every dyad in the *challenges* group was asked, however, to begin with the same challenge, "Where might you find *Prochlorococcus*?" After this first challenge, visitors could select additional challenges. Although they could attempt as many as they wanted, the data collector encouraged every dyad with the *challenges* strategy to try two.

In all three strategies, each dyad was asked to think aloud as they used *Living Liquid* while the data collector took notes on their conversation and their interactions. An interview was administered after the group said they were finished with the exhibit.

Data

The study recruited visitors in December, 2011 and January, 2012. In total, 94 dyads participated in this study, 58 adult-adult, 19 adult-minor, and 16 minor-minor pairs. The interview also asked each dyad if either individual had any background or special interests that helped him or her understand what s/he saw at the exhibit. The tally, according to the three strategies, is shown in Table 1. Chi-square tests did not detect any difference in the distribution of age groups ($\chi 2(2, N = 94) = 3.10, p > 0.05$) or self-reported prior knowledge ($\chi 2(2, N = 94) = 1.28, p > 0.05$) for dyads assigned to the three different strategies.

Table 1
Age group and prior knowledge distribution according to strategy.

Strategy	Group Type			Prior Interest / Knowledge		Total
	Adults	Adult-Child	Child-Child	Yes	No	
Open-Ended	20	4	7	17	14	31
Narrative	19	7	5	15	16	31
Challenge	20	8	4	13	19	32
All	59	19	16	45	49	94

Because *Living Liquid* is a tool, the analysis relied mostly on the think aloud transcripts, which captured thoughts and actions as they happened. Interview data supplemented the think aloud and gave additional insights to what visitors were thinking and their reflections on the overall exhibit experience.

To assess how the three different strategies engaged visitors in data exploration, data coders looked through the transcripts for 1) any mention of each of the data variables encoded in *Living Liquid*, to assess thoroughness of use, 2) any patterns visitors noted including any correlations between plankton population and environmental factors, and 3) counts of visitor-generated questions that were answerable with the dataset with the percentages of those questions that were answered. Approximately 25% of the transcripts were coded by two independent coders to assess inter-rater reliability. The reliability measures are provided in Table 2.

Table 2

Coding scheme used to characterize how visitors explored the data with Living Liquid.

Plankton Type - Visitor mentioned a type of plankton featured in the visualization. This included using its name or noting the color provided that it is clear s/he is using the color to refer to a distinct plankton type.

Code	Example	Inter-rater reliability statistic
Plankton Type: Prochlorococcus	Dyad16: there's green	*kappa* = 0.812
Plankton Type: Synechococcus	Dyad80: little ones! [points to synechococcus]	*kappa* = 0.792
Plankton Type: Dinoflagelletes	Dyad18: dino	*kappa* = 0.697
Plankton Type: Diatom	Dyad112: pink creatures	*kappa* = 0.899

Environmental Variable - Visitor talked about an environmental factor included in the visualization.

Code	Example	Inter-rater reliability statistic
Env Variable: Temperature	Dyad86: cold	*kappa* = 0.899
Env Variable: Sunlight	Dyad27: sunlight	*kappa* = 0.899
Env Variable: Nutrient	Dyad68: nutrients	*kappa* = 0.908
Env Variable: Silica	Dyad2: silica	*kappa* =1

Patterns and Correlation - Visitor commented on a relationship between different types of variables.

Code	Example	Inter-rater reliability statistic
Pattern: Plankton Abundance and Location	Dyad3: Whoa, diatoms! That's a lot. [Looking at Antarctic.]	*kappa* = 0.699
Pattern: Plankton Type and Location	Dyad66: If you are looking for dinoflagellates, you should go here [just south of Australia].	91% agreement [1]
Pattern: Environmental Factor and Location	Dyad86: Nutrients are rich here [Southern Ocean].	*kappa* = 0.732
Correlation: Plankton Type and Environmental Factor	Dyad36: I'm surprised there is so little where it's warm.	*kappa* = 0.817

Question – Visitor asked a question while using *Living Liquid*. Each distinct question was counted.

Code	Example	Inter-rater reliability statistic
Question: Answerable – The question is answerable with the data available in *Living Liquid*	Dyad86: Dino, at the bottom, very small. So should be the same in California?	*ICC* (A,1) = 0.941
Question: Answered – The dyad answered the question while using *Living Liquid*	Dyad86: [Touch spot in Pacific.] Yeah, but we have more dino.	*ICC* (A,1) = 0.948

[1] The kappa value was difficult to interpret because of the very high incidence of this pattern being noted.

Interviews were used to supplement the think aloud data. More specifically, participants' Likert scale rating for how interesting they found the exhibit was compared across the three strategies. In addition, because the *narrative introduction* was designed to give better context for visitors' data exploration, the study looked to see if there was a difference in what visitors thought the point of the exhibit was about for those who watched the slideshow compared to those who did not. Second, the challenges were sequenced so that visitors would have an opportunity to learn that different colors represented different plankton in *Living Liquid*. This mapping between the visuals and referent is fundamental to interpreting the visualization. The analysis, therefore, looked to see if visitors in the *challenges* group were better able to read the visualization.

Results

How do different strategies engage visitors in data exploration?

Thoroughness of use. Two measures were used to gauge thoroughness of use: a) the number of plankton types and b) the number of environmental variables visitors mentioned in their think alouds. In total, the *Living Liquid* prototype showed four plankton types and four environmental variables as listed in Table 2.

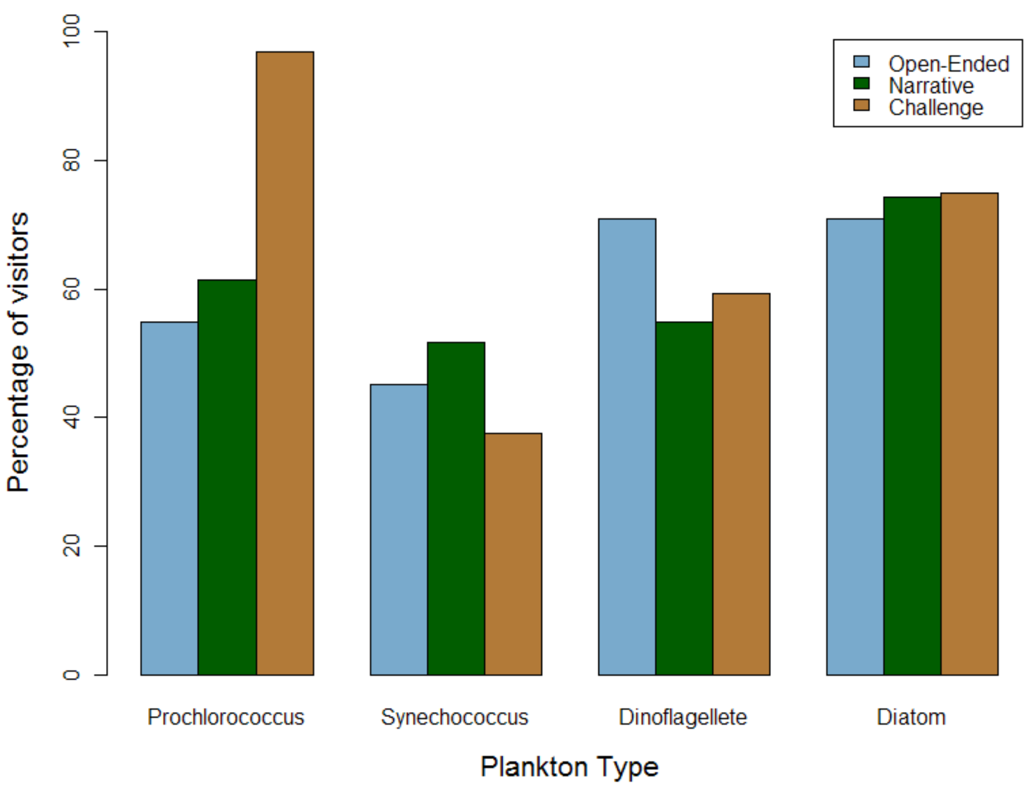


Figure 7. *Percentage of visitors per strategy who mentioned a particular plankton type.*

On average, visitors noticed 3 plankton types. The analysis found no difference among the three strategies. That is; a Kruskal-Wallis test comparing the number of different plankton types visitors noted while using *Living Liquid* detected no statistically significant difference among strategies; $\chi 2(2, N = 94) = 0.568, p > 0.05$.

Figure 7 unpacks the plankton types mentioned by strategy. The percentage of visitors who mentioned *Prochloroccocus* was largest, at 97%, for the *challenges* strategy. This is because the first challenge that visitors were asked to pursue asked them to look for a region on the map that had *Prochloroccocus*. Yet, this initial focus at looking for one particular type of plankton did not lead visitors to look for more plankton types compared to the other two strategies.

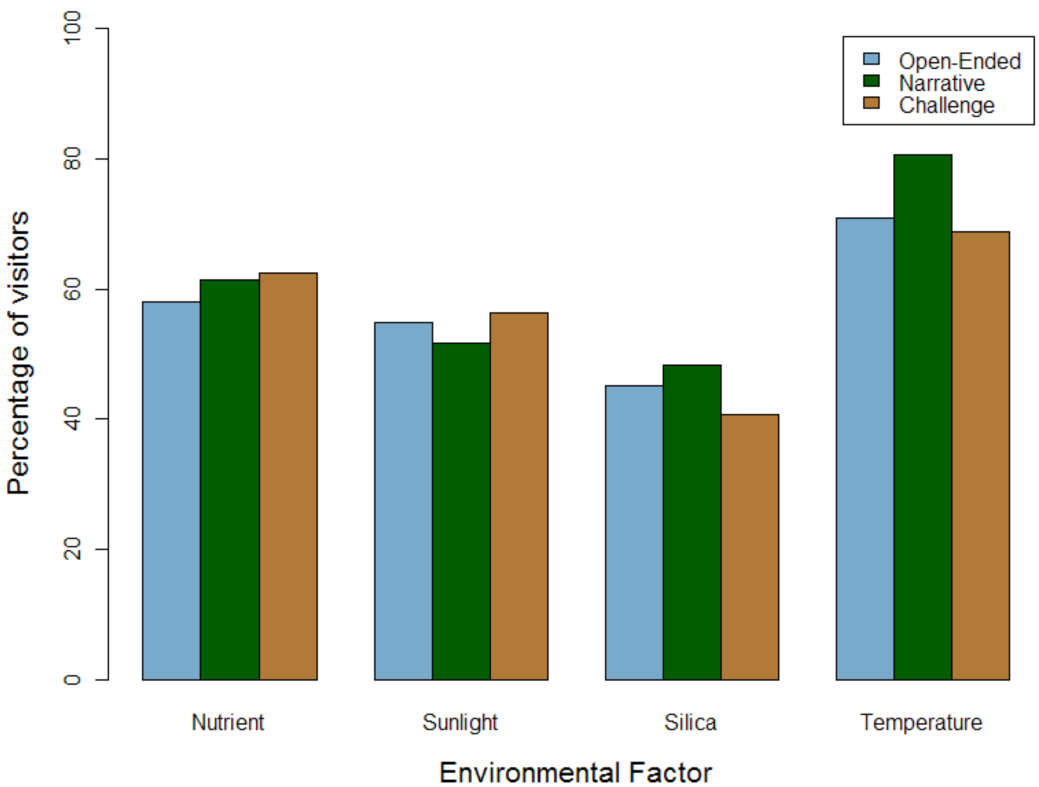


Figure 8. *Percentage of visitors who mentioned a particular environmental factor.*

Visitors, on average, mentioned three environmental factors while using *Living Liquid*. While visitors tended to note temperature most frequently, other less familiar factors, such as silica, were also mentioned. (See Figure 8.) A Kruskal-Wallis test comparing the number of different environmental factors visitors noted found no statistically significant difference among the three strategies; $\chi 2(2, N = 94) = 0.236, p > 0.05$.

The above findings suggest that visitors explored a majority of the data variables visualized in *Living Liquid*, and there was no detectable difference in thoroughness of use among *open-ended exploration*, *narrative introduction*, and *challenges* strategies.

Seeing Patterns. The analysis looked at the percentage of dyads who noted any of the following types of patterns: the geographic distribution of different plankton types, the relative abundance of plankton, the environmental conditions in different locations, or seasonal change in plankton. These are patterns that were encoded in *Living Liquid* in various ways. Colors were used to denote where different types of plankton could be found on the global map, and the circle viewer, which appeared when a visitor touched a particular location on the touchscreen map, showed the icons, in the same color, of the microscopic organisms. The alpha value of the different plankton colors was used to visually encode relative abundance, and the circle viewer showed the relative amount of different plankton types and their overall abundance. Environmental conditions were denoted with level indicators that flanked the circle viewer, and therefore were tied to particular geographic locations. Finally, animation was used to denote change over time with changing color patterns indicating the dynamic nature of the plankton population across the seasons.

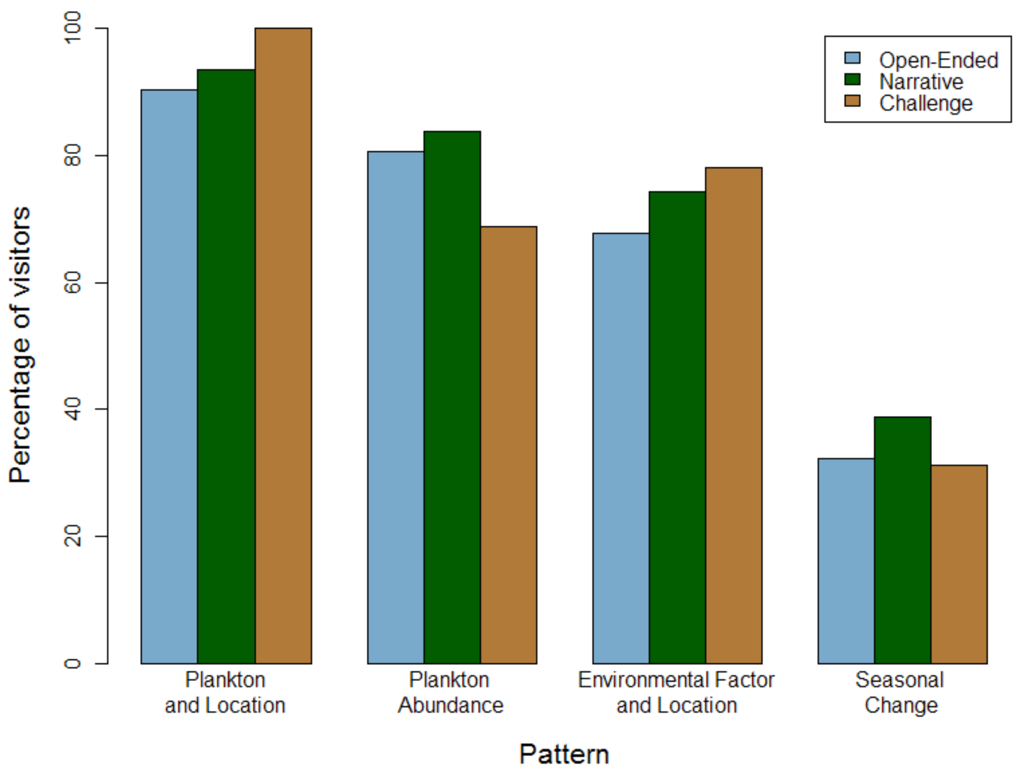


Figure 9. *Percentage of visitors who mentioned seeing a pattern according to the three strategies for engagement.*

Figure 9 shows the distribution of the types of patterns visitors saw during their investigations with *Living Liquid* according to the three strategies. Overall, only a minority (34%) of the visitors made any mention of seasonal change. This suggests that the visualization currently implemented does not make evident how populations are changing over time. These findings echo results from previous studies that revealed difficulties people have in making sense of *dynamic* systems with animation. (See Tversky et al., 2002, for a review.)

A set of Chi-square tests revealed no statistically significant difference among the three strategies in the percentage of visitors who noted any particular type of pattern. The statistics for these Chi-square tests are tabulated in Table 3.

Table 3
Chi-square test statistics for pattern type mentioned for the three strategies.

Pattern Type	***N***	***df***	**$\chi 2$**
Plankton Type and Location	94	2	3.0459
Relative Plankton Abundance	94	2	2.3128
Environmental Factor and Location	94	2	0.8843
Seasonal Change in Plankton	94	2	0.4558

$^*p < .05$

Whereas patterns were encoded in the visualization, correlations depended on visitors seeing relationships that were not explicitly visualized in *Living Liquid*. This study looked for one particular correlation: the relationship between plankton type and environmental factors. This relationship is at the heart of the Darwin Project, which models plankton populations and distributions according to information about the quality and quantity of key environmental parameters for different parts of the world's oceans. Visitors had also asked about this correlation in earlier formative evaluations. Although the key emphasis of the exhibit was on the different plankton at different locations, the analysis looked for this correlation to see if visitors noted this more sophisticated relationship.

Overall, the analysis found that 59% (55/94) of the dyads noticed a correlation between plankton type and the environment. A Chi-square test failed to show any difference among the three strategies. That is, the percentages of visitors who made correlations between plankton and environmental factors did not differ significantly by strategy; $\chi 2(2, N = 94) = 1.28, p > 0.05$.

Asking and Answering Questions. Data coders looked through the think aloud transcripts and counted the number of questions that visitors themselves asked that were answerable with the data visualized in Living Liquid. Note that this count did not include the questions on the challenge cards, which were not authored by visitors. On average, visitors in the *open-ended exploration* and the *narrative introduction* strategy group asked two questions answerable with the data, while visitors in the *challenges* group asked one such question. A Kruskal-Wallis test revealed a significant effect of strategy on the number of questions visitors generated; $\chi 2(2, N = 94) = 7.82, p = 0.020$. A post-hoc test using Exact Wilcoxon Mann-Whitney Rank Sum Tests with Bonferroni correction showed significant differences between the *challenges* ($Mdn = 1$) and *narrative introduction* ($Mdn = 2$) strategies ($z = 2.74, p = .00558, r = 0.346$), and a non-significant trend between *challenges* and *open-ended exploration* ($Mdn = 2$) ($z= 1.86, p = 0.0635, r = 0.346$).

A comparison of the percentage of visitor-generated questions that were answered with the data yielded no detectable statistical difference; Kruskal-Wallis $\chi 2(2, N = 94) = 5.34$, $p > 0.05$. This suggests that challenges did not help better model asking or answering visitors' own questions.

How interesting did visitors find the exhibit?

The results of the Likert test item indicated that visitors overall found the exhibit prototype *Somewhat Interesting* (4), on a scale from *Not Interesting* (1) to *Interesting* (5). There was no statistically significant difference among the three strategies in visitors' interest ratings; Kruskal-Wallis $\chi 2(2, N = 94) = 1.11$, $p > 0.05$. Visitors gave a variety of reasons for why they found the exhibit interesting: 1) they liked seeing patterns in the data; 2) the interactivity and interactions the prototype afforded appealed to them; 3) they learned something they didn't know before; 4) the exhibit was visually appealing; and 5) the subject matter was of personal interest. Many of the explanations visitors provided were not tied to any particular strategy.

Did the narrative introduction better contextualize the exhibit?

The above analyses comparing the three strategies yielded no difference in data exploration for visitors who saw the narrative introduction versus those who had not. Initially, the narrative was designed to situate the visitors' own explorations in the context of the larger research effort to understand global patterns of phytoplankton with data. Was the narrative effective in setting this context? To answer this question, the analysis looked at visitors' responses to a subset of the interview questions they answered immediately after using *Living Liquid*. More specifically, did visitors find personal relevance in the data they explored with *Living Liquid*? Did visitors talk about the broader context when describing the point of the exhibit? Broader context included mention of scientific research about plankton and plankton's role in the larger environment.

Table 4 tallies the number of visitors who self-reported finding something personally relevant in the exhibit's content. A comparison using a Chi-square test between visitors in the *narrative* strategy and the other two strategies found no statistically significant difference in the number who saw personal relevance in the material; $\chi 2(1, N = 94) = 0.548$, $p > 0.05$.

Table 4
Tally of the number of visitors who found personal relevance in the exhibit

Strategy	**Question: While using this exhibit, did you think of anything that related to your own life?**		**Total**
	Yes	No	
Narrative Introduction	12	19	31
Open-Ended Exploration or Challenges	31	32	63
All	43	51	94

Table 5 summarizes the number of visitors who talked about the broader context when asked to describe the point of the exhibit, and gives examples of what counted as 'broader context.' A Chi-square test indicated a statistical difference between visitors who saw the narrative introduction and those who did not; $\chi 2(1, N = 94) = 9.62, p = 0.0019$.

Table 5

Tally of the number of visitors who describe the broader context of the exhibit.

Strategy	**Question: Can you tell me what the exhibit is trying to show overall, if anything?**		**Total**
	Mentioned a broader context	Did NOT mention broader context	
Narrative Introduction	12	19	31
Open-Ended Exploration or Challenges	6	57	63
All	18	76	94
Examples	Dyad28: The importance of plankton and studying it. Dyad48: The need to study more. They are important research. Dyad108: [They] give energy, oxygen. Dyad111: How important plankton are.	Dyad18: Trying to show different patterns of the ocean Dyad56: Types of plankton and seasons and how they change and move. Dyad64: Different plankton types and where they are through the year. Dyad90: It has to do with the ocean.	

Did the challenges better help visitors read the visualization?

The challenge cards were designed to scaffold visitors in exploring the visualization. However, the analyses so far suggest that challenges did not better model question asking and answering behavior. Instead, they were worse at encouraging visitors to ask their own question. At the same time, by first asking visitors to search for one particular plankton type, the *challenges* strategy sought to focus visitors on one variable and its representation and, thereby, scaffold visitors in mapping color to plankton type, which is fundamental to interpreting the visualization. Although they did not encourage visitors to ask their own questions, did challenges help visitors read the data visualized?

The interview data revealed that overall 32% of the dyads had difficulties with the color to plankton type mapping; one or both members of a pair did not know that a particular color was intended to represent a particular plankton type. Table 6 gives the breakdown according to the *challenges* strategy and the other strategies. A Chi-square test revealed that the percentage of the visitors who mapped color to type did not significantly differ according to whether or not they used challenge cards; $\chi 2(1, N = 94) = 0.018, p > 0.05$.

Table 6

Tally of the number of visitors who found personal relevance in the exhibit

Strategy	Thought particular color represented particular plankton type		Total
	Yes	No	
Challenges	22	10	32
Open-Ended Exploration or Narrative Introduction	42	20	62
All	64	30	94

Discussion

This set of comparisons failed to detect any but one difference among the three different strategies in how visitors used *Living Liquid*. There may be a variety of reasons why this may have been so. First, the number of dyads recruited for each strategy was a little over 30, which could detect medium to large effects only, but not more subtle differences. In retrospect, it was then not too surprising to find no statistically significant difference in how visitors engaged with *Living Liquid* particularly between the *open-ended exploration* and the *narrative introduction* strategies. The heart of the visitor experience in both these strategies lied in visitors' using the exhibit itself. The slideshow was but a brief introduction to that experience. Even so, a higher percentage of visitors who did watch the narrative introduction thought that the exhibit was about, in part, the importance of plankton and research on plankton. This indicated that visitors had some awareness of the broader context which the slideshow was design to provide. But, this awareness did not seem to affect how visitors explored the dataset.

Instead, these results indicate that there was no advantage to using a narrative in encouraging data exploration with the visualization tool. This finds some resonance with prior studies on enhancing hands-on exhibits with video narratives. In an earlier study on narratives at interactive exhibits, Allen (2004) found that narratives can add to the exhibit experience by helping visitors make connections between cause and effects illustrated in an exhibit, but they are difficult to create and do not change how visitors use the exhibit for exploration.

In this study, narrative was instantiated as a very short slideshow presented before visualization use. This was but one possible way of incorporating narrative elements into a visualization exhibit. Other visualization tools, such as GapMinder, have interleaved narration with a sequence of data explorations (GapMinder), and narrative elements can be integrated into visualizations in a variety of other ways as well. Recent work by Segel and Heer (2010) is beginning to explicate these different methods. Although the narrative introduction, as implemented in this study, seemed to have no measurable effects on data exploration with *Living Liquid*, we may wish to explore some other narrative devices to not only provide context for the data but also promote exploration and aid in data interpretation.

Whereas the difference between the n*arrative introduction* and the *open-ended exploration* strategies may have been subtle, challenges structured visitors' activities with *Living Liquid*. Nonetheless, the study found that the challenges did not change how visitors explored the data except that visitors tended to generate fewer questions on their own when asked to pursue these staff-authored challenges. If the goal of the visualization tool is to foster question asking, then challenges may actually be impediments to inquiry.

Given the few differences, the findings suggest that efforts in developing visualization tools for data exploration should first focus on tool affordances and integrating interpretative supports for open-ended exploration. Additional questions, such as effective ways of incorporating narrative elements would be worthwhile to pursue to supplement such explorations.

Conclusion

This study contributes to the growing understanding of how to engage the public with exploring scientific datasets in the informal learning context. It points to promising strategies for developing visualization tools for learning in museums: focus on designing affordances in the tool for open-ended exploration; use narrative elements to help visitors see the point of the visualization tool and its data; and consider means other than challenges to encourage question asking. As large-scale data become the currency of many scientific disciplines, a better understanding of how to design visualization tools that allow and encourage the public to explore these data will be a critical part of giving people access to current science as practiced.

References

Allen, S. (2004). *Finding Significance*. San Francisco: Exploratorium.

Allen, S., & Gutwill, J. P. (2009). Creating a Program to Deepen Family Inquiry at Interactive Science Exhibits. *Curator: The Museum Journal, 52*(3), 289-306.

Borun, M., Dritsas, J. I., Johnson, N. E., Peter, K. F., Fadigan, K., Jangaard, A., et al. (1998). *Family Learning in Museums: The PISEC Perspective.* Washington D.C.: Association of Science Technology Centers.

Carroll, J. B. (1993). *Human Cognitive Abilities*. Cambridge: Cambridge University Press.

Edelson, D. C., & Gordin, D. (1998). Visualization for learners: a framework for adapting scientists' tools. *Computers & Geosciences, 24*(7), 607-616.

Frankel, F., & Reid, R. (2008). Big data: Distilling meaning from data. *Nature, 455*(7209), 30.

GapMinder. GapMinder Human Development. GapMinder, 2005. http://www.gapminder.org/downloads/human-development-trends-2005/.

Gutwill, J. P. (2005). Observing APE. In T. Humphrey & J. P. Gutwill (Eds.), *Fostering Active Prolonged Engagement: The Art of Creating APE Exhibits*. San Francisco: Left Coast Press.

Johnson, C., Moorhead, R., Munzner, T., Pfister, H., Rheingans, P., & Yoo, T. S. (2006). *NIH-NSF visualization research challenges report*. Los Alamitos, CA: IEEE Press.

National Research Council. (2009). *Learning Science in Informal Environments: People, Places, and Pursuits*. Washington, DC: National Research Council.

Segel, E., & Heer, J. (2010). Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics, 16*(6), 1139-1148.

Thomas, J., & Cook, K. (2005). *Illuminating the Path: The Research and Development Agenda for Visual Analytics*: National Visualization and Analytics Ctr.

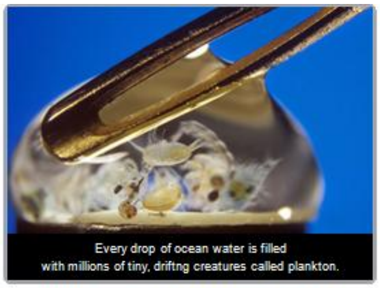
Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: can it facilitate? I*nternational Journal of Human-Computer Studies*, *57*(4), 247-262.

Acknowledgements

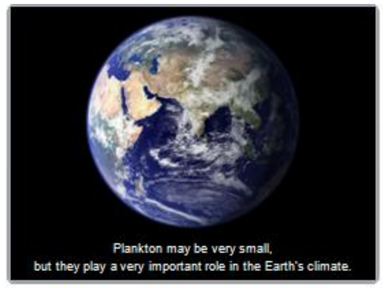
The author would like to thank Mick Follows, Stephanie Dutkiewicz, and Oliver Jahn for sharing the Darwin Project dataset and for being generous with their time and scientific and technical expertise. The *Living Liquid* prototype and its variations used in this study were developed by Jennifer Frazier, Isaac Liao, Eric Socolofsky, and Lisa Sindorf. Ms. Sindorf, Sarah Kimmerle, and Mandy Ice recruited study participants and collected the data. Ms. Sindorf also helped code the data for analysis. This study is based upon work supported by the National Science Foundation (NSF DRL#1011084) and by the Gordon and Betty Moore Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation or the Gordon and Betty Moore Foundation.

APPENDIX A

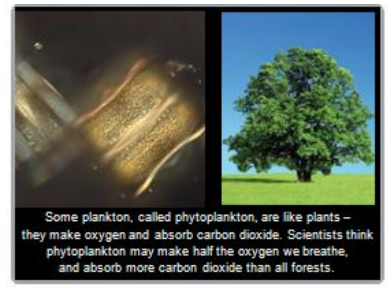
Slideshow



1

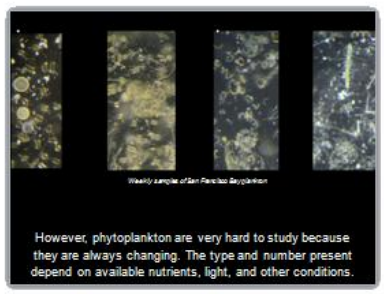


2

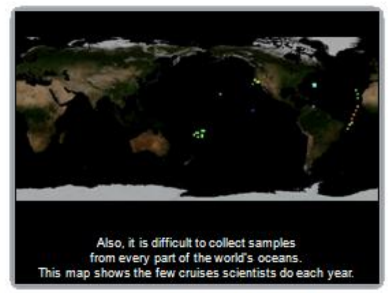


3

4

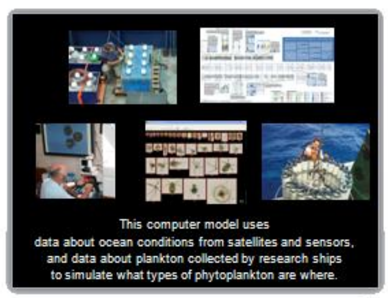


5



6

7



8

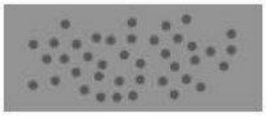
9

APPENDIX B

Challenge Cards

Prochlorococcus Locations

Prochlorococcus are a type of phytoplankton that



- do not need much nutrients.

  (In fact, they can live where no other phytoplankton can survive.)

Where might you find prochlorococcus?

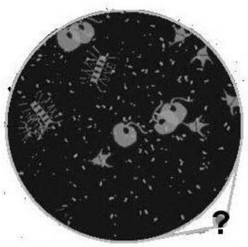
Dinoflagellate Locations

Dinoflagellates are a type of phytoplankton that

- need lots of nutrients
- do not need silica (They do not have hard silica shells.)
- is eaten by tiny fish that are eaten by bigger fish.

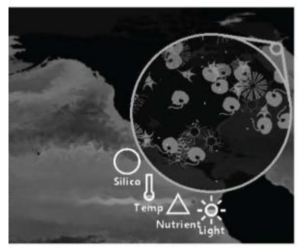
Where might you go to find dinoflagellates (and to fish)?

Phytoplankton Diversity



Where might you find many different kinds of phytoplankton in one place?

**Different Areas
Similar Phytoplankton**



Can you find an area of the ocean that has the same kinds of phytoplankton as the Arctic Ocean?

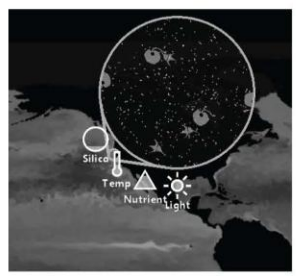
Diatoms Locations

Diatoms are a type of phytoplankton that

- need sunlight to survive.
- need silica to build their shell.

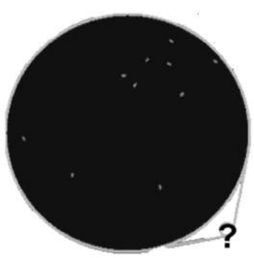
Can you find a warm place where diatoms live?

**Different Areas
Similar Phytoplankton**



Can you find an area in the Atlantic Ocean that has the same kinds of phytoplankton as the waters off the coast of California?

Phytoplankton Abundance



Which part of the ocean has very little phytoplankton?

**Same Area
Different Times
Different Phytoplankton**

Ocean conditions are always changing. The amount of sunlight and nutrients in the water even changes with the time of year. This means that in some areas of the ocean, there can be completely different kinds of phytoplankton at different times of the year.

Can you find a place in the ocean that has completely different kinds of phytoplankton at different times of the year?