Info Vis

Original Article

Information Visualization
1–16

DOI: 10.1177/1473871615594652
ivi.sagepub.com
SAGE

Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors

Katy Börner[1], Adam Maltese[1], Russell Nelson Balliet[1] and Joe Heimlich[2]

Abstract

In the information age, a person's ability to read and make data visualizations is nearly as important as being able to read and write text. This article reports the results of a multi-phase study conducted in informal learning environments in three US science museums. The goal of the study was to determine the familiarity of youth and adult museum visitors with different visualization types. To address this, a total of 273 visitors were shown 5 out of 20 different visualizations that included two charts, five maps, eight graphs, and five network layouts. They were asked to judge the familiarity of the visualization, provide information on how to read it, and provide a name and identify typical locations where they would encounter the data display and possible data sources that might be visualized in this way. The results show that while most participants have a strong interest in science, math, and art, many have a hard time naming and interpreting visualizations. Participants in this study commonly encounter visualizations in school, in books, at work, on the Internet, and in the news. Overall, they were more familiar with basic charts, maps, and graphs, but very few are familiar with network layouts and most have no ability in reading network visualizations. When asked how they would interpret the visualizations, most participants pointed to superficial features such as color, lines, or text as important to developing understanding. Overall, we found that participants were interested in the visualizations we presented to them, but had significant limitations in identifying and understanding them. The results substantiate intuitions shared by many regarding the rather low level of data visualization literacy of general audiences. We hope they will help catalyze novel research on the development of easy-to-use yet effective visualizations with standardized names and guaranteed properties that can be readily used by those interested to understand and solve real-world problems. The results also have implications for how information visualizations are taught and used in formal and informal education, the media, or in different professions.

Keywords

Information visualization, data literacy, user study, science museums

Introduction

Reports such as "Taking Science to School"[1] and "Ready, Set, Science"[2] as well as the new science standards[3] acknowledge that students must understand the nature and development of scientific knowledge and participate productively in scientific practices and discourse. Accordingly, articles in journals on science

[1]Indiana University, Bloomington, IN, USA
[2]COSI, Columbus, OH, USA

Corresponding author:
Katy Börner, Indiana University, CNS, SOIC and IUNI, Bloomington, IN 47405, USA.
Email: katy@indiana.edu

teaching propose methods for data usage in classrooms[4,5] and as a form of evidence.[6,7] Projects such as *Thinking with Data* (TWD) (http://www.rcet.org/twd) conclude there is a critical role for data literacy in providing analytical tools for using data in inquiry, drawing conclusions, and application of data literacy skills across disciplines. TWD points to the benefit of conducting research activities in live learning environments; yet few researchers have investigated how scientists and others read and interpret graphs or understand and use data and data visualization in formal settings,[8–10] and even less on how these skills are cultivated in informal settings. This research is timely as the amount of data in our world is increasing radically, and the capability to analyze datasets is becoming a key basis for all citizens to be data-literate decision-makers.[11]

The National Science Foundation (NSF)-funded *Sense Making of Big Data project* was designed to study how audiences in public spaces relate to and make sense of visual representations of large datasets. The work was conducted by a team of information visualization experts, learning scientists, and museum evaluators that collaborated closely over a 3-year time span. While most user studies in information visualization research use student populations or crowdsourcing, this study interviewed youth and adult science museum visitors across three US science museums. Also, while most studies of visualizations are designed up front and then executed, the study presented here emerged as a series of inquiries which are common in informal science evaluation, that is, each subsequent study builds on questions that emerged as needing greater clarity from the prior study.

The article is organized as follows: the next section discusses related work. It is followed by sections on study design and study results. We conclude with a discussion and outlook.

Related work

Sense-making with data through the process of visualization has been of interest to learning researchers for many years, but most research has looked at individuals applying higher-order thinking to discovering new patterns of communication within a *domain*, which is grounded in a *field*.[12,13] Here, higher-order thinking refers to thinking at synthesis and evaluative levels rather than lower-order thinking such as awareness and understanding. Also of value in sense-making is the capacity for devising data structures and classification models for the sake of knowledge discovery and construction. When given sufficient latitude of control with appropriate tools, children are capable of visualizing sophisticated and complex models of data.[14,15] In addition, a variety of data visualization projects in museums (e.g. *Science on a Sphere* or the *Global Links/Forces of Change* exhibition at the Smithsonian) showed that many visitors were not cognizant of the underlying data[16] or had difficulties interpreting results.[17]

Information visualization researchers frequently use human subject in laboratory studies, longitudinal domain studies, or Mechanical Turk setups.[18] Studies may aim to increase our general understanding of human visual perception,[19] how people read and design visualizations,[20,21] or what makes a visualization memorable.[22] Other studies evaluate the engagement, legibility, memorability, or utility of novel or alternative visualizations[23] or aim to validate the legibility of novel (network) layouts (e.g. can users find certain nodes, edges, paths, clusters, or specific zoom levels?), might involve timed tasks (e.g. finding the shortest path between two nodes and finding the average degree of a set of nodes), exploratory tasks (e.g. what is in the data and where), and targeted tasks (e.g. identify specific nodes that interlink clusters).[24]

An extensive literature review of over 800 visualization publications by Lam et al.[25] identified seven major evaluation scenarios: (1) evaluating visual data analysis and reasoning, (2) evaluating user performance, (3) evaluating user experience, (4) evaluating environments and work practices, (5) evaluating communication through visualization, (6) evaluating visualization algorithms, and (7) evaluating collaborative data analysis. While much research in information visualization focuses on the design and evaluation of interactive visualizations, this study focuses on static visualizations—as shown in the news, textbooks, and so on—exclusively.

It is important to note that many standardized tests exist to assess textual literacy, numeracity, or problem-solving skills, see, for example, the Programme for International Student Assessment (PISA) Assessment and Analytical Framework.[26] However, as of today, no standardized test exists for examining the data visualization literacy of a general population. Promising recent work by Boy et al.[27] proposes a principled way of assessing visualization literacy for bar charts and scatterplots and it might be possible to extend their method to other visualizations such as maps and network layouts.

This study is unique in that it uses a comparatively large number of youth and adult museum visitors. Plus, instead of validating novel visualizations, we are interested to understand whether visitors can read common data visualizations that are frequently used in newspapers, textbooks, or magazines.

Study design

We define data visualization literacy as the ability to make meaning from and interpret patterns, trends, and correlations in visual representations of data. The overarching question for the study was *How familiar are youth and adult visitors with data visualizations?* Specifically, we were interested to understand whether visitors had seen different visualization representations (bar graph, Sankey graph, map of the United States) of different visualization types (chart, graph, map, and network layout) before, where they encounter these visualizations, how they may go about reading different visualizations, how they would call the visualization, and what types of data or information they would visualize in a similar way. The study does not attempt to measure whether visitors can read the data and interpret data visualizations correctly as this is outside the scope possible for a museum floor study. The interview materials and questions were designed by team members who have decades of experience in gathering data in informal learning settings such as museums and they are definitely different from questions that would be asked when interviewing experts.

Visualization framework

Creating easy-to-read but insightful visualizations is difficult as the problem-solving space that needs to be traversed is high dimensional and inherently complex. For example, most data analysis and visualization workflows comprise 10–15 different algorithms (without counting data converter algorithms that ensure the output from the previous algorithms can be read by the next algorithm in line). Most of these algorithms have diverse input parameters, and selecting the best parameter values is non-trivial and impacts visualization results substantially.

Many different taxonomies and frameworks exist on how to structure this complex problem-solving space to make it easier to navigate and manage,[28–38] and some have been implemented in software tools, for example, SYSTAT, SPSS, and OSGI/CIShell macroscope tools (cishell.org).

This study uses the visualization framework introduced in Börner and Polley[28] and Börner.[29] The framework builds on prior works in statistics, information visualization, and graphic design to identify key types involved in the design of insightful visualizations.[39–41] Most relevant for the work presented here, it distinguishes different types of analysis: temporal (when), geospatial (where), topical (what), and network layouts (with whom) and different visualization types: chart, graph, map, and network layout. The latter four are discussed and exemplified here.

Charts. Charts visually depict quantitative and qualitative data without using a well-defined reference system. Examples: pie charts (the sequence of "pie slices" and the overall size of a "pie" are arbitrary; the pie-slice angles and area sizes represent a percentage of the whole) or word clouds (words are randomly positioned; larger words may be set closer to the center to achieve effective use of space or to establish some discernible pattern).

Graphs. A graph maps quantitative and/or qualitative data variables to a well-defined reference system, such as to coordinates on a horizontal or vertical axis. The position of a data point in a coordinate system is determined by the axis values. Examples: timelines, bar graphs, and scatterplots.

Geospatial maps. Maps display data records visually according to their physical (spatial) relationships and show how data are distributed geographically. The positioning of an object on a geospatial map requires a lookup table to convert address data into latitude and longitude information. Examples: world or city maps.

Network layouts. Network graphs use nodes to represent sets of data records and links connecting nodes to represent relationships between those records. Different representations exist for tree and network structures. Node positions might depend on node attributes or node similarity. Examples: *tree graphs*: may be represented as indented lists, dendrograms, node-link trees, circle packings, or treemaps. *Networks*: may be represented by one-dimensional arc graph, tabular matrix diagrams, bimodal network visualizations, axis-based linear network layouts, or force-directed layouts. The first four types use well-defined reference systems (e.g. nodes may be sorted by a node attribute), which means the axes are labeled and their value range is known. Force-directed layouts have no axes.

Note that the different visualization types are preferentially used to answer different types of questions. Timeline graphs are frequently used to answer "When" questions. Geospatial maps are common for answering "Where" questions. Network layouts are excellent for depicting relationships, that is, answering "With Whom" questions.

Study instruments

Using the *Visualization Framework* discussed in the previous section, 20 visualizations were selected from textbooks, news, widely used online visualization

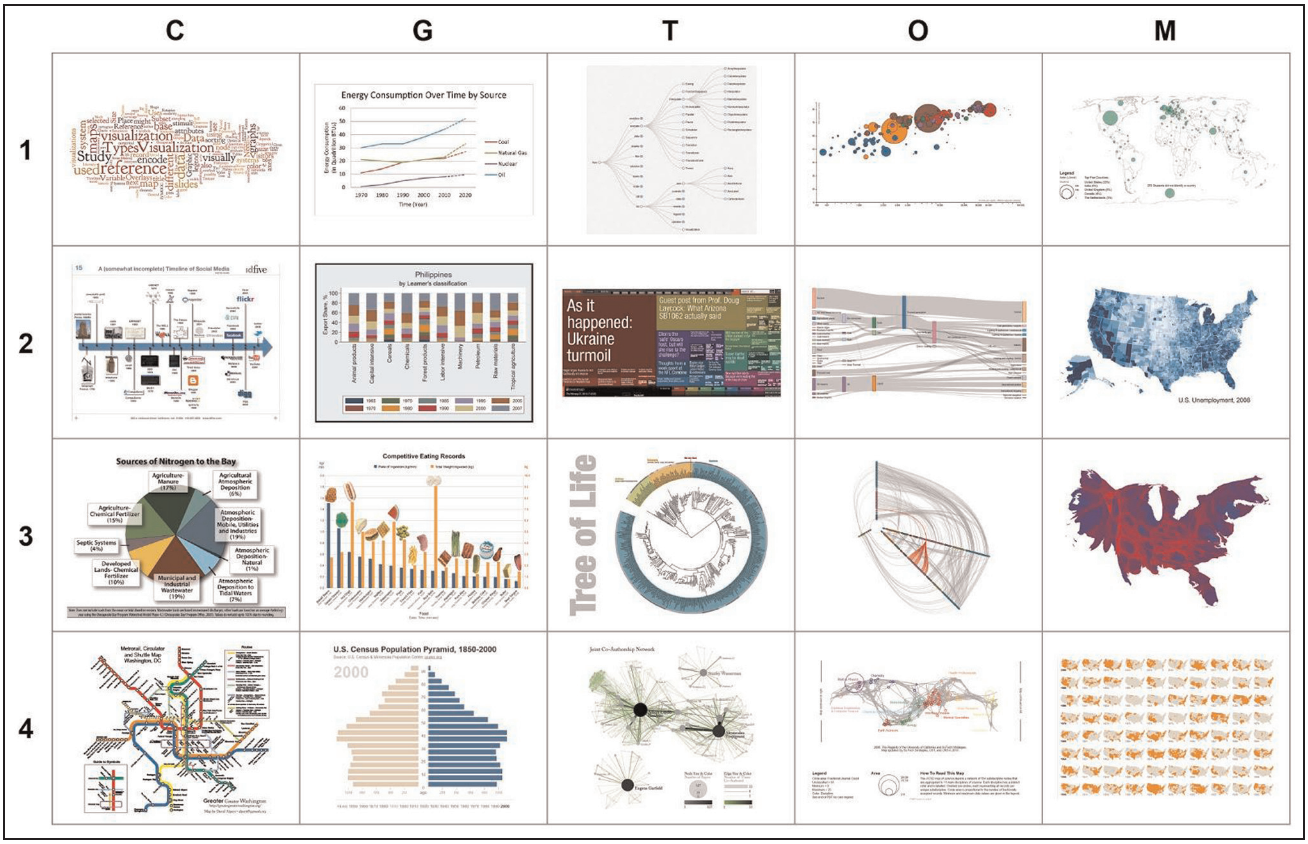


Figure 1. Four sets of five visualizations—each row represents one set; all four rows make up the complete set of all 20 visualizations used in the study. The visualizations are of type chart (C1 and C3), graph (G1, O1, C2, G2, G3, and G4), map (M1, M2, M3, C4, and M4), and network layout (T1, T2, O3, T3, O3, T4, and O4), see Supplementary Materials for high-resolution versions.

libraries such as http://d3js.org, or designed using the Sci2 Tool.[42] Thumbnail versions of all 20 visualizations are given in Figure 1. High-resolution versions or all visualizations are provided in Supplementary Materials. Note that no set of 20 visualizations can possibly cover the rich diversity of existing data visualizations. The final set of 20 does cover all major visualization types discussed in the previous section. Presuming that museum visitors are most familiar with charts and maps, only few examples of these were included, but six graphs and six network layouts are part of the final set. Given that children and adults participated in the study, some visualizations show content relevant for children, for example, the speed eating contest results. Interested to understand the range of visualizations that museum visitors can recognize, some visualizations show more advanced datasets, for example, collaboration networks. When printed in color on letter size paper, all text labels and visual encoding are legible; titles and legends that appear in the original visualization were kept, and none was added.

This set of 20 visualizations was printed in letter size and in color on one printer and laminated so that they had a finished, bright look and would last through the study. The cards were coded on the back by letter for identification and organized into four discrete sets with each set having five visuals (corresponding to the rows in Figure 1, e.g., subject A would see C1, G1, T1, O1 M1; subject B would see C2, G2, T2, O2 M2, etc.) covering different representations. In addition to the 20 visualizations, the other related study instruments used to solicit participants and collect data are included in Supplementary Materials.

Data collection and analysis

Data gathering

Data collection occurred over a 3-week period during spring 2013 at three science museums: The New York Hall of Science (NYSCI) in Queens, NY; the Center of Science and Industry (COSI) in Columbus, OH; and the WonderLab Museum in Bloomington, IN.

Setup. Data were collected within the gallery setting of each institution during normal visiting hours.

Materials. The teams of data collectors at each museum were provided with two sets of the 20 laminated visualizations by COSI. The set 5 of 20 was used in constant rotation; as each museum had two sets of the 20 visualizations, up to two studies could be run in parallel. There were some minor variations in how the visualizations were cycled (i.e. participants saw mixed sets), but these were less than 6% of cases (N = 16) and we do not think this affected the study outcomes.

Procedure. Each site trained multiple data collectors to run the experiment using the *Instructions for Completing the Interview* (see study materials). Since COSI has a relatively high visitation rate, a random sampling method was used to select potential interview participants to ensure a representative population. Sampling was done using focal sampling with a continual ask, a standard practice in museums to avoid selection bias. Focal sampling is creating a visual point across which visitors move. The "ask" is made of the first person to cross that point in the sampling strategy (if a very slow visitation, it may be every third person, though it is impossible to accurately do actual sequential sampling with a random start). A continual ask process has the researcher return to the sampling process upon completion of an interview and ask the first person who crosses the line; this is done continually during the course of the study. The NYSCI and WonderLab venues have less foot traffic which allowed data collectors to approach a greater percentage of visitors ensuring a representative sample of the visitors on the study days. Studies were conducted on both weekdays and weekend days to further represent across audiences.

To initiate data collection, interviewers prompted visitors with variations in the statement: "Would you be interested to participate in a research study?" If visitors refused to participate, this was noted in the *refusal log*. If they gave verbal consent to participate, then the data collector noted group type (single adult, single youth, or youth and adult) and age of visitor(s) on the *data collection forms* (unique case #, location, data collector name, date were prefilled). Then, visitors were asked to state their interest in science, math, and art using a scale of 1 (*not interested at all*) to 10 (*totally love it*). Next, visitors were asked the following five questions for each of the five visuals within a set:

1. Does this type of data presentation look at all familiar?
2. Where might you have seen images like this?
3. How do you think you read this type of data presentation?
4. What would you call this type of data presentation?
5. What types of data or information do you think make the most sense to be included in this type of visual?

Question 1 with a yes/no answer was included to understand whether visitors believed they had ever seen such a visualization. Question 2 was designed to help identify where visitors encounter data visualizations in their daily lives. Question 3 is open-ended but meant to uncover how visitors decode data visualizations. Our pilot studies revealed that visitors use vastly different processes and features to try to make sense of visualizations and they do not have any standard vocabulary to describe how they read visualizations. Related to the notion of relevant vocabulary, Question 4 aimed to collect all names used by a general audience to refer and talk about each of the 20 visualizations shown in Figure 1. Question 5 was included to understand what datasets visitors commonly encounter displayed in visualizations—this also provides important input on what datasets could be used to teach (novel) visualizations.

All data were recorded in the data collection form (see study materials). After that, participants were asked to put the five visuals in order from easiest to most difficult to read. At the end, total time in minutes was recorded and perceived sex was also noted to minimize gender bias in who was asked to participate.

Data preparation

All completed data collection forms from all three venues—195 from COSI, 40 NYSCI, and 38 WonderLab—were scanned and manually transcribed using Microsoft Excel. In some cases, data collectors had to be contacted to discern handwritten notes. In a second round, the Microsoft Excel data were checked for accuracy against the original data collection sheets.

The final sample population comprises 273 visitors: 127 youths ages 8–12 years (with parents/guardians) and 146 adults (over 18 years). Visitors with known perceived sex comprise 110 youth (52% female) and 117 adults (62% female). Within this sample, most individuals saw five visuals, but not every participant answered every question and these blank responses were not analyzed, while "I don't know" responses were; the number of individuals (N) and number of analyzed elements (N^e) were noted in each figure.

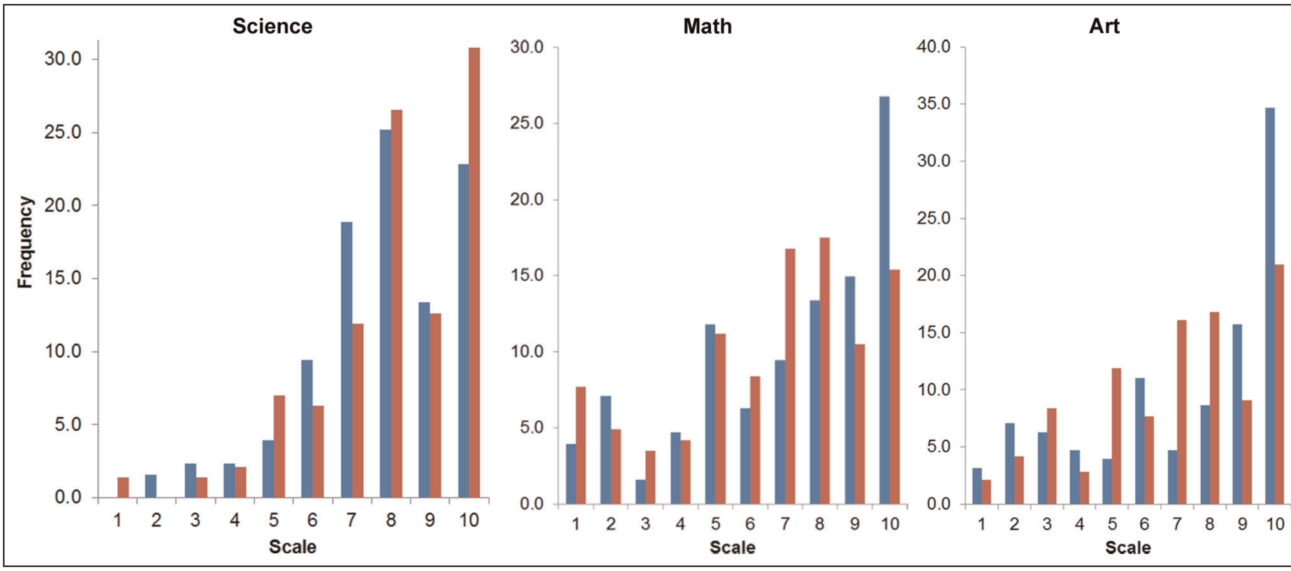


Figure 2. Self-reported interest in science (left), math (middle), and art (right) for 127 youth visitors (in blue) and 143 adult visitors (in red) on a scale of 1 (*not interested at all*) to 10 (*totally love it*).

Study results

To analyze the five questions, a qualitative, iterative, open-ended approach was used in the identification and definition of emergent themes in the data.[43] As expected, each question elicited distinct responses, requiring each prompt to be analyzed independently and leading to the formation of separate coding schemes. We coded at the phrase level, where appropriate, to include any relevant context of a response. Respondents typically mentioned multiple ideas, and thus there are many more instances (i.e. code counts) than participants. All data coding and statistical analysis were performed using the *NVivo 10* software, a platform for analyzing various forms of unstructured data. To better illustrate overall trends in our study, data pertaining to the 20 specific visuals as seen in Figure 1 were aggregated into one of the respective visual types discussed in section "Visualization frameworks" unless otherwise noted.

Pre-questions. Before exploring the set of five visualizations, participants were asked to report their interest in science, math, and art using a scale of 1 (*not interested at all*) to 10 (*totally love it*). Data were collected from 127 youth and 143 adults and were aggregated by relative age and subject, before being binned using the 1–10 scale provided to participants. Any value given outside that range was defaulted to the closest end of the scale. Frequency distributions of interests were calculated for the 127 youth and 143 adults; the results are given in Figure 2. The majority of our sample population is highly interested in science, and less so in math and art.

Five main interview prompts. The results for the five questions asked to the visitors are reported next.

1. Does this type of data presentation look at all familiar?

The yes/no responses to this prompt were cross-tabulated by relative age and visual type.

The results are given in Figure 3. Exactly 62% of the visitors responded that they had seen the visualization before, with the most commonly recognized visual type being charts and the least recognized being network layouts. Age was an important factor as 45% of youths responded that the visuals were unfamiliar to them versus only 31% of adults. Furthermore, adults recognized each of the 20 visuals more often than youths.

2. Where might you have seen images like this?

In coding the *Where* responses, one of the authors initiated coding approximately 20% of the 1000 + responses. The second author coded the same responses using the established codes but adding/editing codes if necessary. After the second pass, the first coder went back and reviewed his coding using the updated codes. At this point, the average inter-rater reliability (measured through Cohen's Kappa) was 0.82. Discrepancies were discussed until agreement

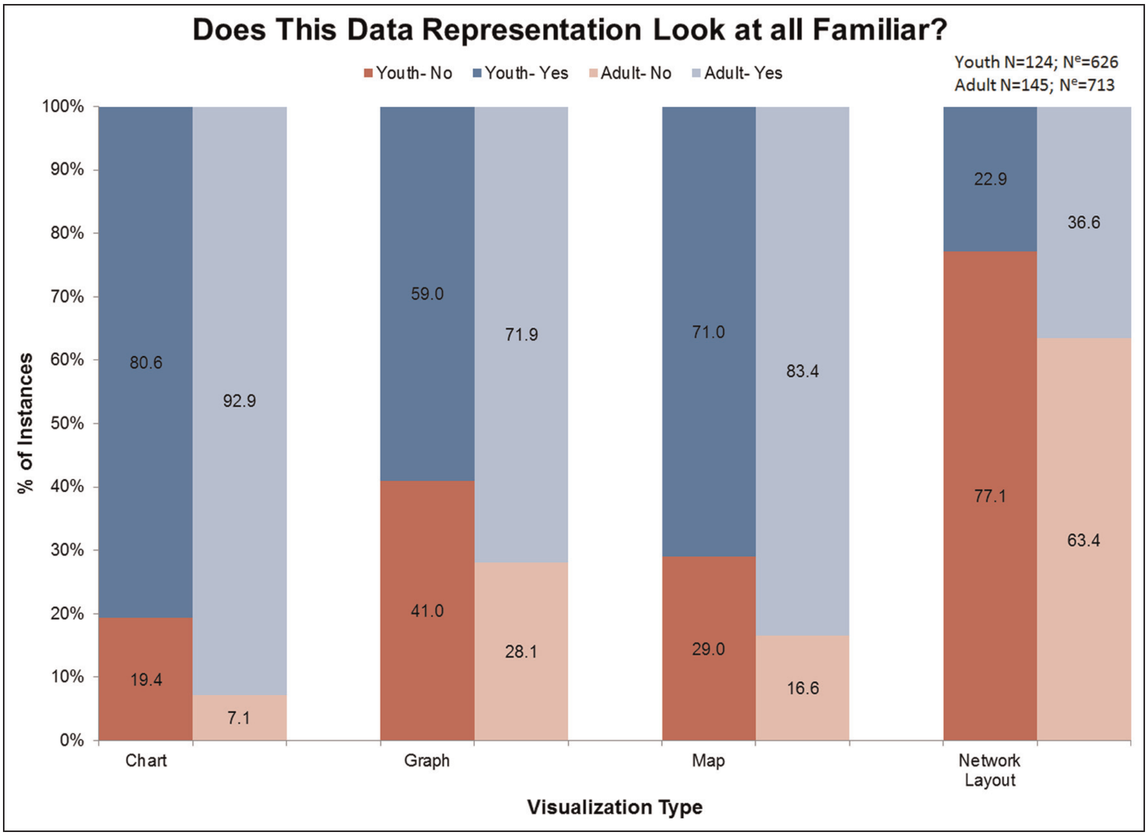


Figure 3. Familiarity of visualizations for youth (blue) and adult (red) visitors.

was reached and codes/definitions were updated. Once this was done, a single coder coded the remaining instances. The codes resulting from this analysis emerged naturally from the data, and as expected, responses were primarily related to physical places (e.g. school, work) or mediums (e.g. newspaper, Internet), giving a sense of where the public encounters visuals.

The results are shown in Table 1. Data were aggregated to higher level categories after extensive discussion; further aggregations (e.g. *news*, *politic*s, and *magazines*) are possible but would further reduce data resolution. Note that some locations are real-world places (e.g. *museums* and *zoos*) while others are online (e.g. *Internet*), or refer to devices (e.g. *technology devices*). Color coding indicates low values (in white) and high values (in dark brown). When asked where they had previously encountered similar visuals, youths reported *school* at a very high rate (52%), with *books* next at approximately 9% of responses. Even then, many of these book references were to textbooks; this is not surprising as a significant portion of a youths' time is spent in school or engaging in school-related activities. *Work and everyday life* was commonly mentioned by adults, but interestingly, adults reported school as the most common place they saw these visuals (20%) suggesting many participants were still in school, referring to their children's schooling, or reflecting on prior experiences. The results also show that in general adults encounter visuals in a wider variety of places more frequently than children and those places tend to be a bigger part of their non-work, non-school activities (news, entertainment, finances, etc.).

3. How do you think you read this type of data presentation?

Answer to this question varied widely. A similar qualitative approach was used with the Q2 response data. As discussed further in the results, the nature of participants' responses led to a more complex coding scheme detailing the key features that participants use to read visuals, what visuals communicate, and the steps they take to understand a visual, among other topics.

The results are given in Table 2. Color coding indicates low values (in white) and high values (in dark green). Participants interpreted this question in several different ways, but responses fell into two main categories: they described components of the visual that participants found useful to reading the visual (*key*

Table 1. Locations where visitors saw the visualizations.

Location Code	Example Response(s)	% of Instances by Relative Age	
		Youth N=123	Adult N=144
Advertisements	Seen it in ads	1.0	2.8
Art	Artwork	0.8	1.1
Books	History Book	8.8	9.7
Data and Research	Research or data that is broken down	0.8	6.3
Entertainment	American Idol	4.3	3.5
Internet	Websites	6.6	8.2
Magazines & Brochures	National Geographic	1.5	5.8
Maps	Anywhere a map would be	7.1	4.4
Medical	Doctor's office, Psychology tests	1.3	1.7
Museums & Zoos	MOMA in NYC	1.8	0.8
News	Newspapers	2.8	10.7
Politics	Voting	0.8	1.7
Posters & Presentations	Business Presentations	0.8	4.2
Public Spaces	I think I've seen this at the mall	5.1	3.8
School	Classes- History	51.8	19.6
Technology Devices	iPod	2.0	1.0
Work & Everyday Life	Weather, At my job in marketing	2.0	12.4
Other	Everywhere	0.5	1.5
Don't Know	Not sure	0.5	0.7
Total Coded Elements		396	710

visual features: 66% of responses) and/or participants described what style of information was being conveyed by the visual (*what the visual communicates*: 14%). Adults tended to describe the trends, relationships, and general information in the visual more often than children (16% and 12%, respectively). Both youths and adults most frequently identified *color* (e.g. "each color represents a line") as a key aspect to understanding the visual (13%), as well as *size and quantity* (10%; e.g. "thickness of lines means how much"), and the presence of *objects* (e.g. circles, pictures, icons: 9%). A small number of responses (n = 23; 78% by adults) noted that certain aspects were missing from the visual and indicated that these pieces would help them better understand the visual (e.g. "there's no key so it makes no sense"). Similarly, many stated that they would use the title, legend, or key to help them read the visual (Table 2) even though those were purposefully not available for them to use (e.g. "guide tells what lines and circles mean"). When discussing what information the visual provided, participants talked about *comparisons or relationships* (e.g. "colors of lines equals different data": 12% of instances), with 62% of these responses provided by adults. Some respondents (7%) indicated that reading a visual was a sequential process and most listed the key pieces with no intended order; most that mentioned this were adults (62%; e.g. "match up things on left and right to see if anything matches—also colors, trace things to middle, continue matching colors in middle with other stuff"). A small percentage of responses provided a specific descriptions of the visual (e.g. "[It's] letting us know where energy comes from": 7%) and responded that they did not know how to read the visual (e.g. "it means nothing": 6%). These results illustrate that people focus on certain aspects of visuals, such as symbols and their characteristics, and that designers should pay particular attention to these when crafting them.

4. What would you call this type of data presentation?

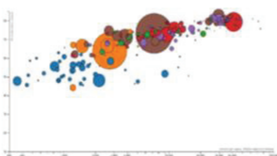
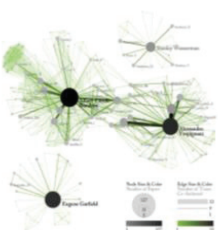
Table 2. Responses for how visitors would read the visualizations.

Response Code	% of Instances by Relative Age	
	Youth N=126	Adult N=146
Identified Key Features	*67.4*	*65.2*
Absence of Information	0.4	1.1
Axes	3.7	3.3
Categories or Groupings	0.6	1.2
Color	13.9	12.0
Key or Legend	6.4	7.0
Lines	3.9	3.8
Location or Orientation of Objects	2.3	2.4
Objects (e.g. circles)	9.8	8.8
Size or Quantity	9.7	11.0
Temporal Aspects (e.g. years, dates, etc.)	4.2	3.3
Title or Labels	3.6	3.1
Generic "Visual" (e.g. "look at the rest of info")	2.6	3.1
Words or Numbers	6.5	5.3
What the Visual Communicates	*12.2*	*16.0*
Comparisons or Relationships	10.4	13.0
General "Information" (e.g. "it breaks things down")	0.8	0.9
Trends or Patterns	1.0	2.1
Other Categories	*20.4*	*18.8*
Procedural (i.e. indicates there are steps to reading visual)	6.6	8.5
Visual Specific (i.e. focused on content of visual)	7.0	6.1
Doesn't Know	6.0	3.3
Blank/No Response	0.8	0.9
Total Coded Elements	*1321*	*1704*

Responses to this question were fit into predetermined categories. Two of the authors independently came up with an "answer key" for each specific visual and then iteratively worked together to create a coding scheme used to categorize responses as being the "technical label," "equivalent phrase," "related phrase," or "unrelated" (see examples in Table 3); a third researcher used this scheme to code all responses, including additional "don't know" and "not applicable" categories. After coding, discrepancies were discussed and resolved through discussion. Table 3 shows four of the visual stimuli used in this study and the terms participants used to name them. These examples illustrate that participants had a wide range of names for the visuals presented to them, and very few would be considered accurate or equivalent terms. This suggests that while participants commonly encounter these visuals in their daily lives, they are unfamiliar with their technical names. Furthermore, the preponderance of different labels the participants assigned to the visuals, even basic ones, suggests that more training and better communication is required before the general public uses a common language to refer to different visualization types.

The results are shown in Figure 4. Participants gave a wide range of answers that were categorized into different levels of relevancy or correctness (see also Table 3). Most respondents had a hard time naming the visuals with any degree of accuracy as very few correctly identified the specific visual (*technical label*: 14%) or came close (*equivalent phrase*: 12%), while an almost equal proportion did not attempt to name the visuals (*doesn't know*: 14%) at all. Some were able to partially identify the visual (*related phrases*: 24%), by either identifying the basic type of visual (e.g. identifying a scatter graph as a graph) or providing a loosely related term (e.g. identifying a scatter graph—also called scatter plot—as a dot chart). The greatest proportion of responses was categorized as *unrelated* (31%); responses that generally made a mindful attempt at naming the visual, but using incorrect visualization terms (e.g. identifying a scatter graph as a map). A very small proportion of responses were categorized as *not applicable* (6%), or responses that were

Table 3. Exemplary visuals and example responses for different levels of relevance/correctness.

Example visuals	Relevancy categories with examples					
	Vis type	Technical label	Equivalent phrase	Related phrases	Unrelated	Not applic.	Doesn't know
C1 	Chart	Word cloud Tag cloud	Word chart Wordle	Info graphic Visual thesaurus Word cluster	Messaging Advertising Randomness	Mess Test Makes me think of words	Not sure No idea
O1 	Graph	Scatter graph with proportional symbol coding	Scatterplot Bubble graph Rainbow scatter graph Bubble chart Weighted scatter plot graph	Word scramble Population graph Graph to show income and life Crumb graph Dot chart Graph	Mixed up graph Paint dots Countries of the world Open map Circley graph Chart	Things that have color Messy	No idea Not sure
M2 	Map	Choropleth map	Map of United States Density distribution map Density map Map with concentrations	Map Population chart Population map Percentage map Saturation concentration Hue map	Graph Chart	Boring A chicken	Not sure No idea
T4 	Network layout	Force-directed layout	Relationships graph	Connection chart Networking map Web graph	Spider web diagram Spider web population Web charts	Thingy Spider web Line graph	I don't know No idea

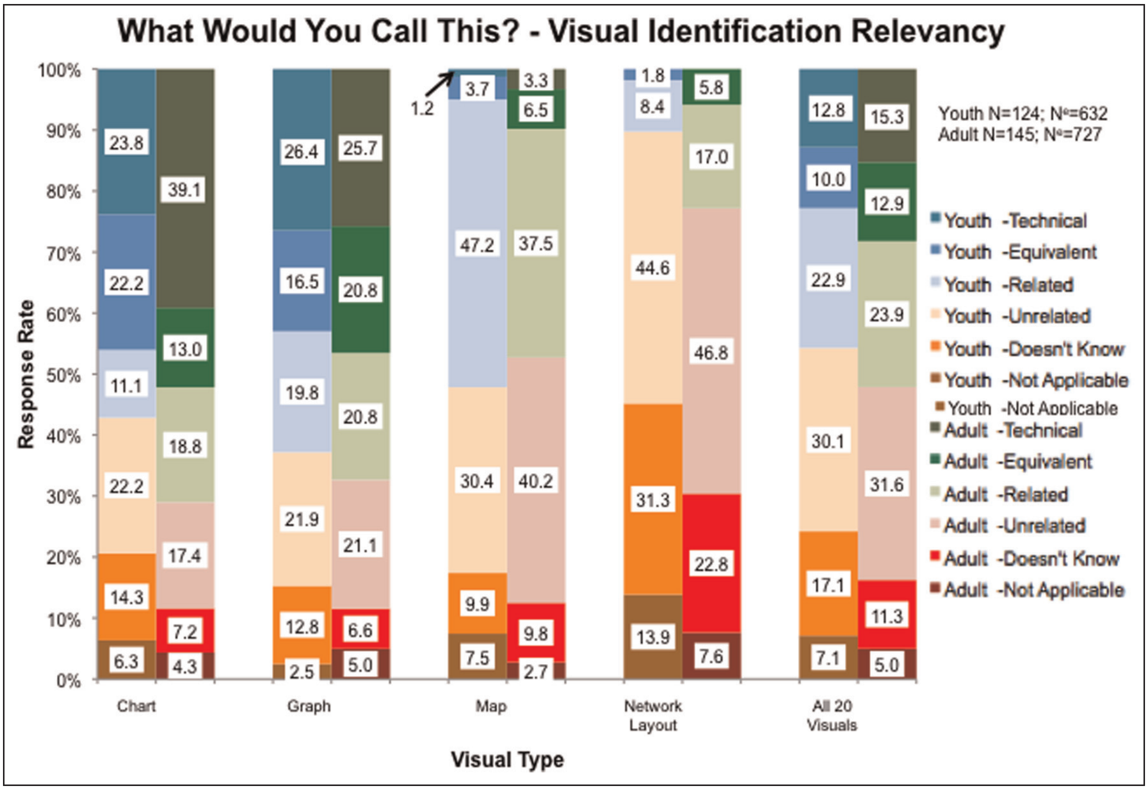


Figure 4. Names youth and adult visitors gave the visualizations by different levels of relevance/correctness.

completely unrelated to visualizations (e.g. "a chicken"). In total, less than 50% of respondents gave answers that would be considered at least partially accurate. Note that while some readers might think that calling T4 a *spider web diagram* is very creative, we are interested to understand whether visitors are using the correct terms to refer to visualizations. If different names are used, for example, *spider web diagram* and *connection chart*, then two visitors may never realize they talk about the very same visualization.

Specifically, network layout visuals were the least frequently identified visuals by participants. The five different network layouts were seen a total of 337 times and no participants fully and correctly named them; only 4% and 13% of responses to these visualizations were categorized as equivalent or related, respectively. The vast majority of responses were incorrect (46%) or non-attempts at identification (i.e. *doesn't know*: 27%), meaning that over 83% of responses failed to identify the visual to any degree. Of all the visuals, maps had the highest proportion of *related* responses (42%) and a high proportion of unrelated responses (36%), indicating that participants were able to generally identify the visual as a map, but could not use the specific term (2%) or something equivalent (5%). The most recognizable visual types were charts and graphs, with 64% and 65% or responses indicating some level of correct identification, but charts were more frequently given a technically correct identification (32%) than graphs (26%) were. Within the graph category, not all visuals were identified with the same degree of success. Specific graphs such as the timeline graph (85%), bar graph (66%), and line graph (59%) were frequently given a technically correct label, while other graphs, such as the Sankey diagram and treemap, had a relatively high frequency of *unrelated* responses (45% and 55%, respectively). Treemap visualizations are a special case as the technically correct name includes the name of the wrong visualization type: map.

A second, broader qualitative analysis was performed to see how often participants mentioned *chart*, *map*, *graph*, or *network layout* (see section "Visualization framework"). Responses were coded for mentions of these terms regardless of their "correctness" or surrounding words and the usage of the word "network" was categorized as network layout.

The results are given in Table 4. Color coding indicates low values (in white) and high values (in dark green). Almost half of responses (44% of responses across all visuals) did not include any of the visualization keywords chart, map, graph, or network layout to name the visuals. Of the responses that did, the word graph occurred most often (48% of responses using a

Table 4. Occurrences of *chart*, *graph*, *map*, or *network layout* in the names visitors gave the visualizations.

	What would you call this?			
Visualization Type (# Used in Study)	*Chart*	*Graph*	*Map*	*Network Layout*
Chart (2)	63.5	31.1	5.4	0.0
Graph (8)	20.1	75.4	4.4	0.0
Map (5)	10.8	16.9	71.9	0.4
Network Layout (5)	36.0	40.4	21.1	2.6
All (20)	23.9	48.0	27.6	0.5

keyword); it should be noted that there were eight graph visuals used in the study likely leading to the higher usage of this word. Even though there were only two charts used in this study, 24% of responses referenced the word chart in some way. The other two visualization types, maps and network layouts, both had five visuals in the study and their associated keywords were mentioned 28% and 0.5% of all responses.

Charts were most commonly associated with the word chart (64%), but also frequently mislabeled as a graph (31%), and occasionally the word map was used (5%). Maps were not strongly misidentified as any other visual type in particular, with the words chart (17%) and graph (11%) mentioned most often. Graphs were the most highly recognized visual with 75% of responses using the word graph in their response. Graphs were most often misidentified as a chart (20% of the time), and less frequently as a map (4%); the most commonly misidentified graphs were the Sankey diagram (chart: 37%) and treemap (map: 45%). As previously discussed, participants were generally able to identify maps (72%), but participants were unable to provide further details. As suggested above, network layouts were the most unfamiliar type of visual to the participants. Network layouts were the only group of visuals to be labeled a different visual more often than the correct visual, with the phrase network layout occurring in only 3% of responses. The word graph was used most often (40%), with chart being used almost as frequently (36%); they were misidentified as maps 21% of the time. The high amount of cross-identification between charts and graphs indicates that the general public uses these terms interchangeably, and in some cases as a general purpose term for all visuals.

5. What types of data or information do you think make the most sense to be included in this type of visual?

When participants were responding to a more familiar visual, they were able to provide responses that indicated they understood the primary purpose of the visual type or they responded with a novel and appropriate use for that visual. For example, when shown a timeline of social media innovations, many individuals were able to describe its purpose (e.g. "development of something over time," "things that have an order like time," "dates and years and events") or provide alternate content ("American history. Drug response to medication. History development."). Less familiar visuals tended to produce "mimicked," vague, or "don't know" responses. For example, network layouts prompted many individuals to simply read back information from the visual they were looking at (e.g. "number of times someone coauthored or influenced other authors"), provide a generalized response ("things from science class," "different types of chart"), or no response at all ("I don't know").

Post-questions. A total of 53 subjects sorted the five visuals in order from easiest to most difficult to read. We asked them to do this so that we could get a sense of how they compared the visuals in terms of difficulty. The values were calculated across rows to illustrate the ranked position that participants most frequently placed the visual types.

The results are given in Table 5. As can be seen, charts are judged easiest to read, followed by maps, and then graphs. Networks are singled out as hard to read. Charts were most commonly placed in the easiest to read section. Graphs and maps had similar placement frequencies, and network layouts were overall ranked hardest to read.

Study limitations

How much time visitors spent with the visualization had a major impact on their ability to accurately

Table 5. Sorting results for chart, graph, map, or network layout.

	Participant Rankings (N=53)				
Visualization Type	*1 (Easiest)*	*2*	*3*	*4*	*5 (Hardest)*
Chart	34.8	4.3	26.1	17.4	17.4
Graph	25.5	24.5	17.9	18.9	13.2
Map	21.7	31.9	21.7	17.4	7.2
Network Layout	4.5	6.0	19.4	25.4	44.8

answer the five questions. People did spent more time if the data shown in the visualization were of interest to them, for example, children were interested to explore the food eating contest results. Other people were drawn to visualizations due to color or overall aesthetics, and sometimes that translated into a willingness to spend time interpreting the visual, but other times remained as a cursory, purely aesthetic appeal.

The visualization framework used in this study to judge the correctness of one of the five questions: "What would you call this type of data presentation?" was developed over the last 15 years. The framework draws on leading works by major visualization pioneers, is one of the most comprehensive frameworks in existence today, and it has been used to teach Information Visualization at Indiana University for more than 10 years and for the last 3 years in the Information Visualization MOOC (IVMOOC) that students from 100 + countries attend. However, it is not agreed upon by all visualization experts, that is, not all expert would agree with the four visualization types and their description or the names it gives to specific visualizations. In fact, alternative names used in the press and scholarly papers, for example, metro map or treemap, make it hard to identify the correct visualization type, for example, network layout, not map. Expert agreement on standard naming conventions and major visualization types would help avoid confusion and support the transfer of visualization sense-making skills within and across visualization types. Clearly, names for specific visualizations and visualization frameworks will need to co-evolve to make them maximally useful.

Discussion and outlook

Collectively, the results reported here provide strong empirical evidence that a very high proportion of the studied population, both adult and youth, cannot name or interpret data visualizations beyond very basic reference systems, that is, they have rather low performance on key aspects of data visualization literacy. The results confirm statements by leading visualization practitioners arguing that most US citizens cannot read the visualizations that are common in newspapers, textbooks, or on the web.

The results from this particular study provide important "ground truth" about the non-existence of a visualization language (i.e. an internally consistent set of proper names for different visualization (types) that are well defined and consistently used) and the data visualization literacy of museum visitors. This parallels challenges identified for "big data" research, namely, the need to "help remove the barrier associated with different uses of terminology." A recent report[44] argues that because of "the interdisciplinary nature of Big Data, there are many concepts and terms that are defined differently by people with different backgrounds" and suggests short courses and webinars to identify and help resolve differences.

It should be pointed out that language—human's ability to acquire and use complex systems of communication—is a key component of human development. Without language, that is, the ability to name and label physical or imaginary objects and concepts, it is much harder if not impossible to learn or talk about them.[45] While visualization experts might be comfortable with ambiguity and the nature of subtle differences in names for visualizations, non-experts would much benefit from proper names and definitions for common visualizations that can be learned once and effectively used over a lifetime.

Taking practitioner comments and empirical evidence seriously has a number of serious implications for information visualization education and research. In the information age, the ability to read and make data visualizations is increasingly important. Given this importance and the low skill level, it is highly desirable to teach data visualization reading and writing skills in formal and informal education settings. As for information visualization research, much effort is focused on the development of new, ever more complex visualizations—few of them are ever used outside of information visualization conference proceedings or journals. However, if youth and adults with a predilection toward science—apparent by the fact that they visit science museums—cannot read simple visualizations, more work should focus on empowering many to read and understand basic visualizations. Close collaborations with educators, journalists, government, and industry seem highly desirable to improve our understanding of how to make, use, and teach data visualizations more effectively.

Although our results suggest that charts and graphs are the most commonly encountered visuals in school or work, more research needs to be done to explore the amount of visualization literacy individuals gain in school, especially with more complex maps and network layouts. Recently, select elementary and middle schools have started to teach data visualizations—a unique opportunity for visualization researchers to study and help improve the acquisition of data visualization literacy. Plus, the use of data is now a key component of recent efforts to create national standards in mathematics[46] and science;[3] there is a unique opportunity for visualization researchers to study and help improve the acquisition of data visualization literacy.

The general public is likely to encounter new and complex visuals outside of formal learning settings putting more of an onus not only on their creators to provide guidance but also on educators to teach students broad and explicit visualization reading strategies so that they are better equipped to understand novel visualizations they encounter. Being able to name visualizations correctly will empower different stakeholders to refer to and discuss a diverse range of visualizations. Being able to see a new visualization and identify its type makes it possible to transfer knowledge (e.g. on how to read a graph) from known visualizations to the new visualization.

Other studies conducted in this project (all reports available on informalscience.org) show that construction of complex visualizations led to more accurate interpretation than deconstruction. There were also suggestions in the findings related to interest: interest was found to vary by age and sex in subjects, in topics of visuals, and in the appeal of the visuals. Yet, there were no topics that were consistently appealing across respondents. The same was found in the appeal of the visual itself.

Going forward, we are interested to advance this research by collaborating with others to conduct similar investigations in different environments and countries and with alternative visualizations. Toward this end, all study materials have been made available online at http://cns.iu.edu/2015-VisLit.html. Please feel free to contact the authors for more information.

Acknowledgements

The authors would like to thank data collectors at the WonderLab, COSI, and NYSCI. The Sci2 Tool has been used to generate several of the visualization stimuli; it uses the Cyberinfrastructure Shell (http://cishell.org) developed at the Cyberinfrastructure for Network Science Center (http://cns.iu.edu) at Indiana University. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Funding

This work was carried out through support from the National Science Foundation under DRL award number 1223698.

Supplementary materials

The study materials for *Recognition and Meaning Making in Data Representations* available at http://cns.iu.edu/2015-VisLit.html include the following:

- Data collection basics;
- Instructions for completing the interview;
- Data collection form;
- 20 Visual stimuli;
- Refusal log;
- Blank data entry spreadsheet.

References

1. National Research Council. *Taking science to school: learning and teaching science in grades K-8*. Washington, DC: The National Academies Press, 2007.
2. Michaels S, Shouse AW and Schweingruber HA. *Ready, set, science! Putting research to work in K-8 classrooms*. Washington, DC: The National Academies Press, 2008.
3. NGSS Lead States. *Next generation science standards: for states, by states*. Washington, DC: The National Academies Press, 2013.
4. Eckhaus A and Wolfe R. Gathering and interpreting data: an interdisciplinary approach. *Sci Scope* 1997; 21(4): 44–46.
5. Espinoza F. Developing inquiry through activities that integrate fieldwork and microcomputer-based technology. *Sci Activit* 2002; 39(3): 9–17.
6. Hug Barbara and McNeill Katherine L. Use of first-hand and second-hand data in science: does data type influence classroom conversations? *Int J Sci Educ* 2008; 30(13): 1725–1751.
7. McNeill KL and Krajcik J. Middle school students' use of appropriate and inappropriate evidence in writing

scientific explanations. In: Lovett MC and Shah P (eds) *Thinking with data: proceedings of the 33rd Carnegie symposium on cognition*. Mahwah, NJ: Lawrence Erlbaum Associates, 2007, pp. 233–265.

8. Kozhevnikov M, Motes MA and Hegarty M. Spatial visualization in physics problem solving. *Cognitive Sci* 2007; 31(4): 549–579.
9. Maltese AV, Harsh JA and Svetina D. Interpretation of graphical representations along the novice—expert continuum. *J Coll Sci Teach*, in press.
10. Roth W-M and Michael Bowen G. Professionals read graphs: a semiotic analysis. *J Res Math Educ* 2001; 32(2): 159–194.
11. Manyika J, Chui M, Brown B, et al. *Big data: the next frontier for innovation, competition, and productivity* (Report for the McKinsey Global Institute). New York: McKinsey & Company, 2011.
12. Csikszenymihalyi M. *Creativity: flow and the psychology of discovery and invention*. New York: Harper Perennial, 1996,
13. Sternberg RJ and Lubart TI. Chapter 1. The concept of creativity: prospects and paradigms. In: Sternberg RJ (ed.) *Handbook of creativity*. New York: Cambridge University Press, 1999, pp. 3–15.
14. Hammerman JK and Rubin A. Strategies for managing statistical complexity with new software tools. *Stat Educ Res J* 2004; 3(2): 17–41.
15. Konold C. Handling complexity in the design of educational software tools. In: *Proceedings of the seventh international conference on teaching statistics*, Salvador, Brazil, 2–6 JULY 2006. Auckland, New Zealand: IASE.
16. Storksdieck MK, Bronnenkant C, Kissler L, et al. *The national museum of natural history's global links—forces of change: summative exhibition evaluation*. Annapolis, MD: Institute for Learning Innovation, 2003.
17. Grack N, Ellenbogen K, Rowe S, et al. Beyond Oohs and Aahs: using evaluation to help visitors understand complex scientific visualizations. *Paper presented at the visitor studies association conference*, Grand Rapids, MI, 25–29 July 2006.
18. Kittur A, Chi EH and Suh B. Crowdsourcing user studies with mechanical turk. In: *Proceedings of the SIGCHI conference on human factors in computing systems (CHI)*, Florence, 5–10 April 2008, pp. 453–456. New York: ACM.
19. Heer J and Bostock M. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In: *Proceedings of ACM human factors in computing systems (CHI)*, Atlanta, GA, 10–15 April 2010, pp. 203–212. New York: ACM.
20. Grammel L, Tory M and Storey M. How information visualization novices construct visualizations. *IEEE T Vis Comput Gr* 2010; 16(6): 943–952.
21. Yi JS, Kang Y, Stasko JT, et al. Understanding and characterizing insights: how do people gain insights using information visualization? In: *Proceedings of the 2008 workshop on beyond time and errors: novel evaluation methods for information visualization*, Florence, 5 April 2008, pp. 4:1–4:6. New York: ACM.
22. Borkin MA, Vo AA, Bylinskii Z, et al. What makes a visualization memorable? *IEEE T Vis Comput Gr* 2013; 19(12): 2306–2315.
23. Skupin A, Biberstine JR and Börner K. Visualizing the topical structure of the medical sciences: a self-organizing map approach. *PLoS ONE* 2013; 8(3): e58779.
24. Bahador S, Simonetto P, Kobourov S, et al. Node, node-link, and node-link-group diagrams: an evaluation. *IEEE T Vis Comput Gr* 2014; 20(12): 2231–2240.
25. Lam H, Bertini E, Isenberg P, et al. Empirical studies in information visualization: seven scenarios. *IEEE T Vis Comput Gr* 2012; 18(9): 1520–1536.
26. OECD. *PISA 2012 Assessment and Analytical Framework* (Technical report). OECD, 2012, http://www.oecd.org/pisa/pisaproducts/PISA%202012%20framework%20e-book_final.pdf
27. Boy J, Rensink R, Bertini E, et al. A principled way of assessing visualization literacy. *IEEE T Vis Comput Gr* 2014; 20(12): 1963–1972.
28. Börner K and Polley T. *Visual insights: a practical guide to making sense of data*. Cambridge, MA: The MIT Press, 2014.
29. Börner K. *Atlas of knowledge: anyone can map*. Cambridge, MA: The MIT Press, 2015.
30. Card SK and Mackinlay JD. The structure of the information visualization design space. In: *Proceedings of the IEEE symposium on information visualization*, Phoenix, AZ, 21 October 1997, pp. 92–99. Los Alamitos, CA: IEEE Computer Society.
31. Chi EH. A taxonomy of visualization techniques using the data state reference model. In: *Proceedings of the IEEE symposium on information visualization*, Salt Lake City, UT, 9–10 October 2000, pp. 69–75. Los Alamitos, CA: IEEE Computer Society.
32. Keim DA. Visual exploration of large data sets. *Commun ACM* 2001; 44(8): 38–44.
33. Kosslyn SM. Understanding charts and graphs. *Appl Cognitive Psych* 1989; 3(3): 185–225.
34. Mackinlay JD. Automating the design of graphical presentations of relational information. *ACM T Graphic* 1986; 5(2): 110–141.
35. Munzner T. *Information visualization: principles, techniques, and practice*. Natick, MA: A K Peters, 2014.
36. Pfitzner D, Hobbs V and Powers D. A unified taxonomic framework for information visualization. In: *Proceedings of the Asia-Pacific symposium on information visualisation*, Adelaide, Australia, vol. 24, Feburary 2003, pp. 57–66. Darlinghurst, NSW, Australia: Australian Computer Society, Inc.
37. Shneiderman B. The eyes have it: a task by data type taxonomy for information visualizations. In: *Proceedings of the IEEE symposium on visual languages*, Boulder, CO, 3–6 September 1996, pp. 336–343. Los Alamitos, CA: IEEE Computer Society.
38. Wilkinson L. *The grammar of graphics*. New York: Springer, 2005.
39. Bertin J. *Semiology of graphics*. Madison, WI: University of Wisconsin Press, 1983.

40. Engelhardt Y. *The language of graphics: a framework for the analysis of syntax and meaning in maps, charts, and diagrams*. PhD Dissertation, University of Amsterdam, Amsterdam, 2002.
41. Harris RL. *Information graphics: a comprehensive illustrated reference*. New York: Oxford University Press, 1990.
42. Sci2 Team. *Science of science (Sci2) tool*. Indiana University and SciTech Strategies, 2009, https://sci2.cns.iu.edu
43. Glaser, BG. The Constant Comparative Method of Qualitative Analysis. *Social Problems* 1965; 12(4): 436–445.
44. AAPOR Big Data Task Force. AAPOR report on Big Data, 2015, http://www.aapor.org/AAPORKentico/AAPOR_Main/media/Task-Force-Reports/BigDataTaskForceReport_FINAL_2_12_15.pdf
45. Stork S and Sanders SW. You say potato, I say potäto: problems associated with the lack of shared meaning in instruction and learning. *Quest* 2000; 52(1): 60–78.
46. National Governors Association Center for Best Practices, Council of Chief State School Officers. *Common core state standards*. Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010.