ARTICLE

Begin at the Beginning: A Constructionist Model for Interpreting Data Visualizations

MARY ANN WOJTON, DONNELLEY HAYDE, JOE E. HEIMLICH, AND KATY BÖRNER

Abstract Within museums and science centers, the visual presentation of data represents a timely, relevant, and formidable interpretive challenge. Tackling this challenge head on, however, means employing a series of principles that position educators to support learners of all skill levels in interpreting data visualizations more generally. In this article, we introduce the Simplicity-Familiarity Matrix, a research-driven model for situating complex data visualizations in the context of exhibition design. This model emerges from a study of data literacy that was undertaken at five informal learning institutions, along with established principles of constructionist approaches to teaching. Specifically, it highlights key affordances and challenges we associate with data visualizations along two spectra: simplicity and familiarity. We propose the Simplicity-Familiarity Matrix, along with criteria for each quadrant, to assist museum professionals when designing interpretative materials for an exhibition space. In light of these considerations, we close with several guiding principles for supporting learners' apprehension of data visualizations in museums: (1) including the use of well-designed, easy-to-understand data visualizations, (2) providing appropriate support and guidance, (3) offering multiple modes of access, and (4) making data visualizations relevant to visitors, e.g., via personalization.

INTRODUCTION

Data visualizations are becoming ubiquitous, and making sense out of data representations is necessary in order to understand and begin to utilize big data in interpretive decision making. Although constructivist approaches are critical to acknowledging learners as whole people and supporting individual meaning-making, the work of interpreting data visualizations to public audiences presents specific tactical challenges. We seek to address these challenges with a set of guiding principles for working with complex data visualizations in informal learning settings. We also introduce the Simplicity-Familiarity Matrix, a model which can support concrete decision-making about how to conceptually situate and incorporate data visualizations in exhibition design. The resources were developed in response to major takeaways from Indiana University's *Sense-Making of Big Data* project, a federally-funded study of how learners apprehend and perceive complex data visualizations (an element of NSF grant #1223698), with additional data collected in the context of the Aquarium of the Pacific's *Our Instrumented Earth* data education project (NASA grant #NNX12AL16G*)*. These findings emerged from learners' responses to large, complex data sets, and therefore they are particularly well-situated for the range of topics and skill levels addressed in exhibition settings. Since we began

Mary Ann Wojton (mwojton@cosi.org) and Donnelley Hayde (dhayde@cosi.org) are Research Associates with COSI's Lifelong Learning Group. They study learning in informal settings including museums and zoos. Joe E. Heimlich (jheimlich@cosi.org) is Co-Director of COSI's Lifelong Learning Group; his research focus is lifelong learning about and in the environment, with interests in integration of social role, context, and conditions of the visit. Katy Börner (katy@indiana.edu) is the Victor H. Yngve Distinguished Professor of Engineering and Information Science in the School of Informatics, Computing, and Engineering, Founding Director of the Cyberinfrastructure for Neetwork Science Center at Indiana University, Bloomington, IN.

with particularly difficult interpretive cases, the research-driven model that we offer may support informal learning professionals in interpreting nearly any data visualization.

Why Big Data?

Across the project research that underpins this writing, "big data" are considered to be data sets that are representative of large demographic groups, those that are exhaustive (or nearly exhaustive) for their topic or niche, and those that provide the basis for longitudinal study (Börner et al. 2015). Any presentation of big data can be broken down into constituent parts: its reference system (i.e., visualization type such as scatterplot, map, or network), its data overlay (i.e., geometric symbols such as point, line, text, symbols), and its visual encoding (i.e., graphic variables like position, form, color, and texture of geometric symbols (Börner and Polley 2014; Börner 2015).

Because big data sets are increasingly accessible and public discourse about data has become a fixture of daily life, timely concerns for educators and other interpretive professions include how to expand learners' data literacy, as well as how data relate to learners' real-world experiences. A major affordance of large, complex data sets is that they allow investigators to see patterns, trends, and structure across vast numbers of cases. Therefore, effective visualization consists of "aggregating purely local relationships in a large dataset, to build a simplified picture of its global structure" (Wolfe 2013).

Meanwhile, research and evaluation continue to yield data about how learners process information and about how informal learning spaces can support authentic, meaningful learning. Simplification, focus on key patterns, and thoughtful presentation of new information are critical to supporting learners' understanding of big data because complexity and volume increase a viewer's cognitive load, or how much of an individual's working memory is required to process the visuzalization (Krumhansl et al., 2012, p. 20). Although big data sets are common in informal learning settings that represent scientific research (e.g., science centers, aquariums, zoos, etc.), the representation of large, complex data is also germane to the work of other museums. In helping visitors make meaning of spatial, temporal, and social relationships between collections objects, historical events, artistic movements, and/or their own experiences, most institutions find themselves tasked with showcasing data in a way that makes sense to people, especially in the form of maps, timelines, and diagrams. This paper aims to support professionals in this work by offering a bridge between established theory and practical application in museums and other informal learning settings.

Big Data: Large data sets, those that are exhaustive (or nearly exhaustive) for their topic or niche, that may be analyzed computationally to reveal patterns, trends, and associations.

Data Visualization: An image that is representative of the raw quantitative or qualitative data and supports exploration, examination, and communication of the data, or the process of creating a data visualization (cf. Azzam et al. 2013).

Visitors as Learners

Existing research on the processes and habits of mind that relate to learners' apprehension of data visualization strongly reflects a shared theory of teaching and learning that will likely be familiar to many educators, curators, exhibition designers, and other interpretation professionals. For example, Thudt et al. (2012) recommend including elements of collaboration and play, as does Canning (2013). This supports the well-accepted principle within museum learning that in order to effectively implement constructivist learning experiences, visitors must both have the agency to explore topics according to their interests and have appropriate levels of support when they are challenged by new or unfamiliar ideas (cf. Hein 1998; Hohenstein and Moussouri 2018).

Meanwhile, scholars have also observed a need for varying depths of exploration; interpretation should support learners whether they are interested in top-level highlights or detailed, comprehensive information (Hinrichs et al. 2008a,b; Thudt et al. 2012; Ma et al. 2012; Canning 2013; Krumhansl et al. 2012). Others emphasize simplicity of presentation, noting that data visualization should not include extraneous information (Kostakos et al. 2013; Kaminski and Sloutsky 2013; Krumhansl et al., 2012). Within the field of formal learning, scholars have observed that effective teaching related to data visualization seems to involve both introducing conventions of systems of representation and, to an even greater extent, illustrating how those conventions of systems relate to real-world phenomena. (Bright and Friel 1998; Roth and Bowen 1994). Studies related to learners' engagement with data visualization have suggested that teachers can be most effective by helping students focus on managing complexity in order to help them move from abstract representation to big-picture, concrete meaning (Konold and Higgins 2003).

Theories of attention-perception and cognitive schema suggest that people enter informal learning contexts holding a set of prior experiences, beliefs, values, and ideas – all of which shape their expectations and the things to which they give attention. In general, people seek out information and experiences that support held beliefs and values, and they often do not notice or attend to experiences and information that do not fit with their entrance narratives (Doering and Pekarik 1996). In short, visitors generally have their attention focused on the experience of the learning context itself – in this case, a science center or museum – rather on specific target learning content (i.e., a complex data visualization). Further, visitors enter informal learning spaces with a range of literacies and awareness of specific topical content. For example, many visitors have not heard the term "big data" or may not understand what it involves, and museum visitors' knowledge of and familiarity with data visualizations also vary widely (Tranby et al. 2013; Wojton et al. 2014). An especially critical point from previous literature is that visitors in any informal learning setting should be able to approach and understand interpretation, whether they already know a lot about a given topic or if they are encountering the topic for the first time (Hinrichs et al. 2008a,b; Ma et al. 2012; Hachet et al. 2013). This idea underpins the major premise of a constructionist approach.

Process as a Strategy

Constructivist learning frameworks following Vygotsky and Piaget (e.g., Hein, Rogoff, and others) have widely influenced interpretation in modern museums, science centers, aquariums, and zoos. While these philosophical approaches have helped informal learning professionals acknowledge and value individual learners' meaning-making, there is less support for the *process* of making meaning as it applies to complex learning tasks. In response, the model we suggest draws on a constructionist theory of learning (cf. Kim 2001; Kafai and Resnick 1996).

Following Papert and Harel (1991), we observe the distinction between *constructivism* as an epistemology and *constructionist* strategies situated within that epistemology. Whereas constructivism acknowledges learners as individuals who make meaning from the world based on their existing knowledge and their personal lenses of experience and culture, constructionism represents the possibility for facilitating meaning-making for those individuals through active learning and the process of building a shareable result. To accomplish this central task of constructionism, a critical feature of the teaching approach is interpreting from the ground up. In short, although learners are assumed to have existing knowledge, experience, and perspective that influence their meaning-making, there is no assumption from the outset about *what* these dimensions include; rather, meaning is built (Papert and Harel 1991; Kim 2001).

By starting from basic concepts and building to more advanced ideas, learners may move at their own pace and explore to the level of depth and detail that best suits their interests and needs. For our purposes, using a constructionist lens to interpret visual representations of big data would mean providing access points through which learners could begin to sequentially and gradually grasp key findings communicated through elements such as format, color, and number. While this approach remains consistent with a shared constructivist philosophy, we aim to suggest concrete tactics that can support the more granular concerns of interpretive practice.

As described below, the construction element of the *Sense Making of Big Data* project essentially used the process of building a data visualization to explore what specific features of data visualizations are easier and harder for visitors to understand accurately and quickly. In doing so, it demonstrated that helping learners both arrive at and move beyond key points of recognition seems to support the strongest understanding of intended meaning for the greatest number of people. Thus, we anticipate that being able to better articulate where a given visualization falls in terms of its relative complexity and familiarity to museum visitors will help museum professionals identify elements of any given data visualization that may require more interpretive support. At the same time, building on common entry points within data visualizations may reveal pathways toward providing visitors with engaging, interesting opportunities to apply their own existing knowledge to the interpretive experience.

Meanwhile, we as researchers have also adopted a constructivist lens (Lincoln and Guba 2013) in the process of articulating this model, and we recognize this lens as central to effectively applying the model, too. Drawing on Lincoln and Guba (2013), we understand our work as dialogic and reflexive: to arrive at the premises and recommendations described below, we have attempted to consider our own positionality, the knowledge base of our assumed readers, and the ways in which our data do – and do not – reflect a construction of audience knowledge that can be reasonably compared to what we assume as “average” education about and experience with data visualization. Furthermore, we acknowledge that the most fruitful uses of the Simplicity – Familiarity Matrix will be those that consider practitioners’ own understandings of their audience and context, then invite museum visitors to co-create meaning in a process of sense-making that reflects a shared logic.

> **Constructivism:** A learning theory in which learners make meaning from the world based on their existing knowledge and their personal lenses of experience and culture,
>
> **Constructionism:** A constructivist pedagogical approach that involves creating opportunities for all learners to build an object that is shareable. Although constructionist interventions acknowledge that learners bring a range of prior knowledge to the experience, the activity itself begins at a shared starting point.

METHODS

The *Sense Making of Big Data* project was designed to study how audiences in public spaces, in this case those in informal learning contexts, relate to and make sense of representations of large data sets. The study was guided by the research question, how do people engage with/understand reference systems? Four visuals were used, combining math, science, history, and art to convey four diverse topics. Data were collected at four museums and one aquarium: COSI in Columbus, Ohio, the Marian Koshland Science Museum in Washington, DC, the New York Hall of Science in Corona, New York, WonderLab Museum in Bloomington, Indiana. Also included in this discussion are data collected at the Aquarium of the Pacific in Long Beach, California, which replicated the study as an interactive data activity at a NASA Night event as part of its *Our Instrumented Earth* project. A total of 250 adult and 138 youth participated in this study.

Adults, alone or with youth, met with the data collector on the floor of the participating museum or aquarium. The data collector had four different visualization booklets composed of the layers necessary to create the data visualization. These spiral bound booklets consisted of a base printed in color on white cardstock and laminated, and layers printed in color on transparencies. The booklets were coded on the back by letter for identification. The visualization booklets were used in constant rotation. During construction, the individual/group was shown the base and asked to describe what it was communicating. The evaluator noted what they heard and then asked the individual or group to describe what the base and layer 1 was communicating, then the base, layer 1 and 2, then the base, layer 1, 2 and 3, etc. until they had completed all the layers of the visualization.

STUDY FINDINGS AND IMPLICATIONS

Findings from this study suggest two major dimensions of data visualizations themselves: simplicity and familiarity. Each is explored in more detail below.

Data Visualizations Exist on a Continuum From Simple to Complex

A key insight that emerged from this study, as well as other studies from the *Big Data* project, is the relationship between the complexity of a data visualization and guests' ability to make meaning from those representations. In short, study findings suggest that the visiting public has a relatively low level of visualization literacy (Börner et al. 2016). Meanwhile, data visualizations seem to be situated on a continuum of increasing complexity, where poles can be represented as "simple" and "complex." We define "simple" as data visualizations that involve dichotomous (two categories) data, such as pictographs, tables, pie charts, and bar charts – those that are typically taught in elementary school

(Alper et al. 2017). Other simple data visualizations may utilize common reference systems, such as a map of the United States combined with simple graphics (lines and circles) and labels.

Data visualizations become increasingly complex with the addition of new dimensions or variables – for example, when additional geometric symbols are added, or multiple, possibly animated graphic variables (color, texture, motion) are used. While visitors may have encountered the basic format of a data visualization through school or the mainstream media, they may not understand a third axis on the X-Y graph or the difference between stacked bars on a stacked bar chart. Complex visualizations, such as a three-dimensional scatter plot and heat map, are likely to involve more features than visitors typically see and/or results from calculations and analyses that are not usually introduced until high school and college math and science classes (e.g., higher level algebra and trigonometric functions).

Complex data visualizations that represent three-dimensional objects may enhance an exhibit aesthetically and intellectually. However, these represent another level of complexity as two-dimensional representations are limited to a single plane. Therefore, more complex visuals may require increased spatial abilities among visitors in order to understand another level of abstraction. This may require additional scaffolding, provided by an educator or additional explanatory text. For example, if a history museum believes the topography of a battle site may aid a visitor in understanding the outcome of a battle, they could display a topographic map of the battle site. To increase visitor understanding, they could add appropriate interpretive supports, including an explanation of the contour lines and the geometric and pictorial symbols that represent various terrain, hydrographic, and transportation features relevant to the battle.

Data Lie on Continuum From Familiar to Unfamiliar

Separate from the visualization format, the data themselves also affect visitors' facility of meaning-making. This study found that a museum visitor's knowledge of and experience with the data displayed within the visualization exists on a separate continuum anchored by unfamiliar and familiar. Familiar data are found in common media or as part of the everyday lexicon; examples might include sports statistics and weather reports. In contrast, unfamiliar data would include information that is not part of everyday life, be in limited use in specialized fields, such as maps used by biologists and wildlife specialists to show spatially-referenced animal tracks or genetic maps used by parents to understand their unborn child's chance for inheriting a disease.

SIMPLICITY-FAMILIARITY MATRIX

Overlaying the two continua creates a four quadrant graphic (see Figure 1) that museum professionals may find useful when designing interpretive materials as a platform for educational intervention. In short, considering the familiarity of data in conversation with the complexity of the representational system being used to describe it may reveal strategies for constructing visualizations with appropriate, sequenced scaffolding for visitors to informal spaces. Quadrant A includes data visualizations that are simple and consist of familiar data; B includes more complex visualizations with familiar data; C includes simple data visualizations with unfamiliar data; and D consists of complex data visualizations with unfamiliar data. Criteria for each quadrant follow.

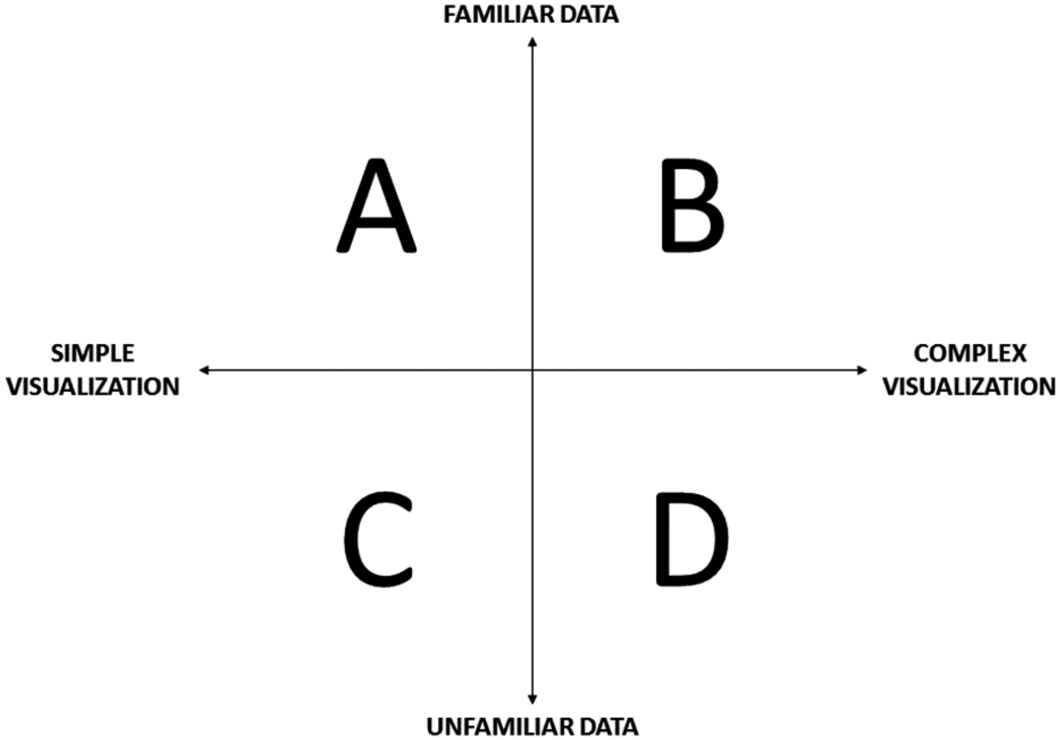

Figure 1. Data visualization quadrant map.

Quadrant A: Simple and Familiar

A visualization from this study that would fall in Quadrant A is *Airports Reachable from Chicago O'Hare International Airport in 2008* (Figure 2). This was a familiar graphic for the majority of participants; over 80% of the participants in this study were able to understand this visualization in the fastest average time (4 minutes). It combines a common reference system (the map of the United States) with simple graphics (lines and circles) and labels that were familiar to the majority of participants, i.e. locations of major airports and number of flights per day. Other visualizations that would be placed into Quadrant A might include weather maps, charts and graphs like those typically found in the USA Today Newspaper, and XY line graphs such as those found in elementary school texts.

Quadrant B: Complex and Familiar

Quadrant B includes complex visualizations with familiar data. For example, the *Gapminder World 2012: Wealth and Health of Nations* data visualization (Figure 3) would fall into this quadrant. It is a modified XY graphic where lines are replaced by bubbles of different sizes to depict different relationships which moves it toward the more complex end of the visualization continuum. However, in this study, 85% of the individuals and groups who constructed this graph understood the content because "health" and "wealth," words which appeared on the base layer, are quantifiable

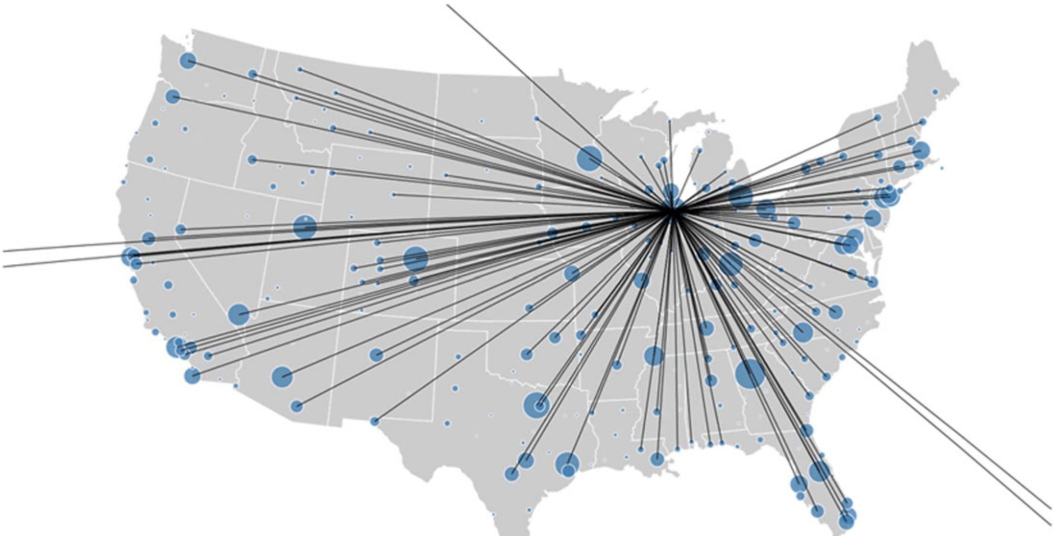

Figure 2. Airports reachable from Chicago O'Hare International Airport in 2008. [Color figure can be viewed at wileyonlinelibrary.com]

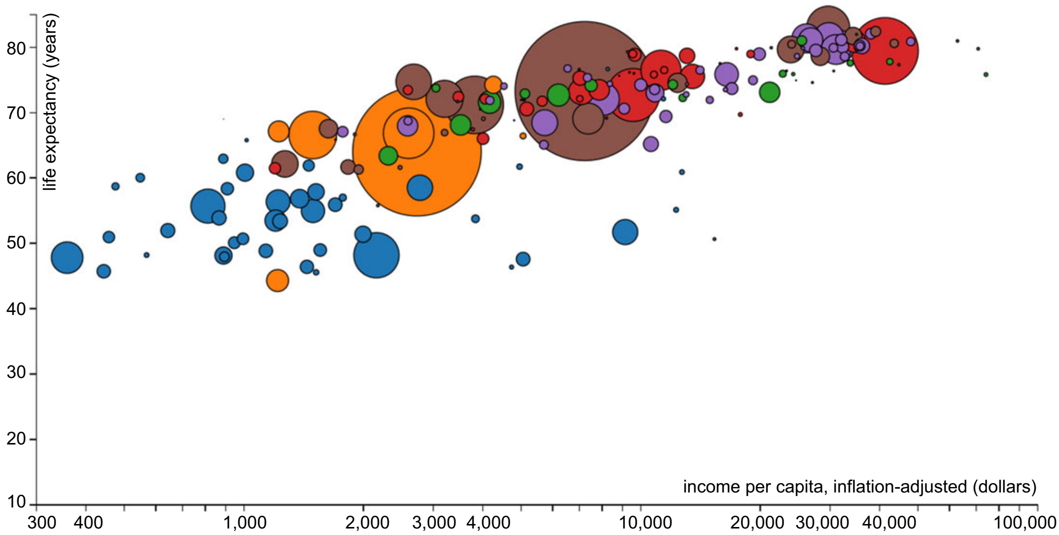

Figure 3. Gapminder World 2012: Health and Wealth of Nations. [Color figure can be viewed at wileyonlinelibrary.com]

words that were familiar to the majority of the participants in this study. Other types of visualizations that might be included in Quadrant B are temperature and location parallel coordinate plots and stock market information displayed on stream graphs.

Quadrant C: Simple and Unfamiliar

Quadrant C includes simple visualizations with unfamiliar data. In this study, the *Competitive Eating Records* graphic would be in Quadrant C. Although the double bar graph is a step

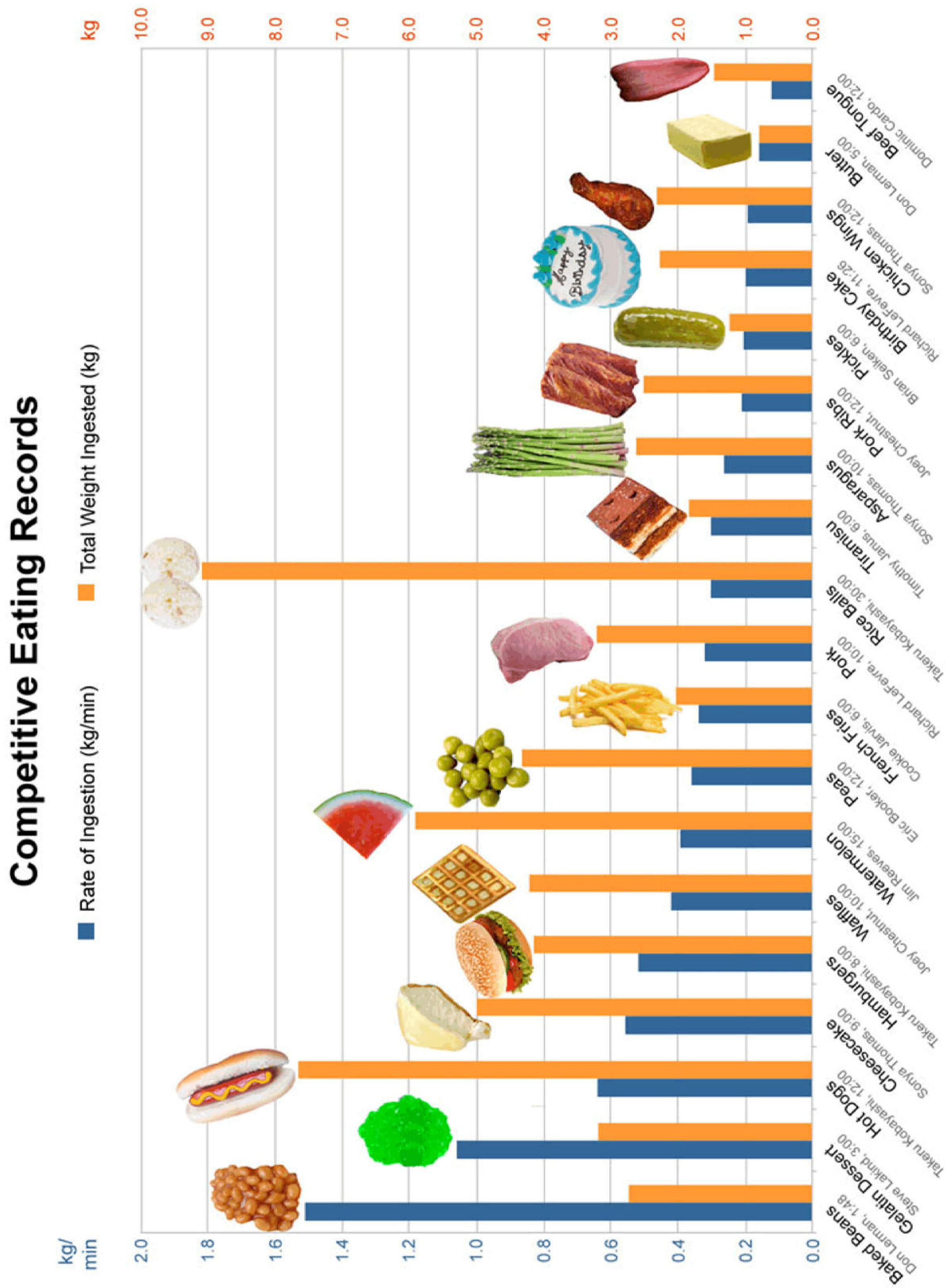

Figure 4. Competitive eating records. [Color figure can be viewed at wileyonlinelibrary.com]

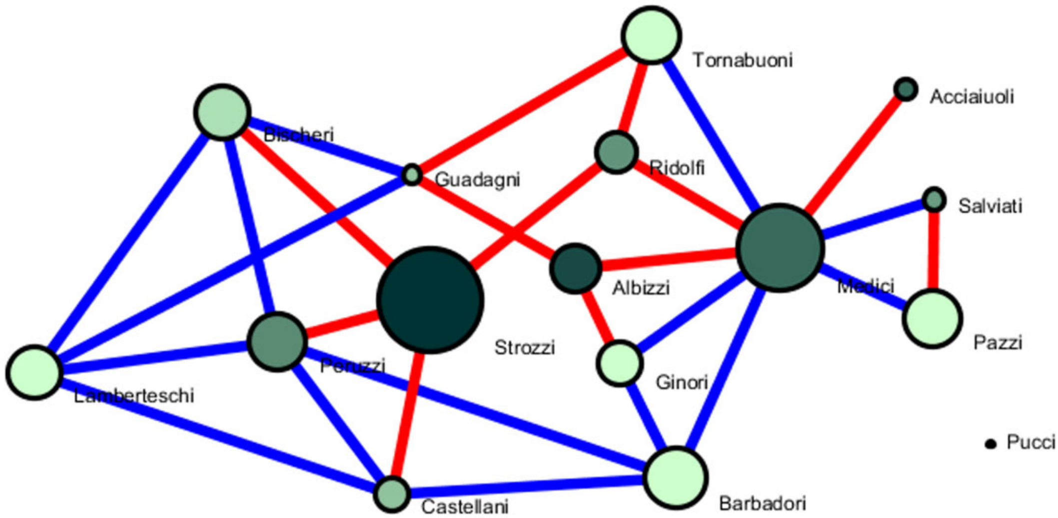

Figure 5. Padgett's Florentine families. [Color figure can be viewed at wileyonlinelibrary.com]

above a simple bar graph, the presentation of this data was familiar to most, however only 64% understood the visualization represented competitive eating records. The majority of participants understood this graphic dealt with food, but unless they recognized names of the competitive eaters, or understood the title, they did not completely understand this visualization (See Figure 4).

Quadrant D: Complex and Unfamiliar

Quadrant D includes complex visualizations with unfamiliar data. In this study, a network map of Florentine families would be placed in Quadrant D. *Padgett's Florentine Families* (Figure 5) used a complex reference system (network map) to present information that was unfamiliar to science museum and aquarium visitors, possibly because it referenced Italian history, which was not a common frame of reference for respondents at science museum and aquariums. On average, participants constructing this data visualization took over 2 minutes longer to make meaning of this graphic than participants who constructed the simple and familiar data visualization (6:06 minutes compared to 4 minutes). Even after reading the title participants' level of understanding was not always evident. After seeing the complete graphic there were participants that indicated they still didn't know what it meant.

Rather than suggesting that museums simplify data visualizations so they are less intimidating for those with a lower comfort level with quantitative information (Korn 2015), we suggest museums accept that some visitors have challenges making meaning of complex data visualizations and provide appropriate interpretive support for data visualizations that are complex or include unfamiliar data.

THEORY INTO PRACTICE

To determine how to place a data visualization within the Simplicity-Familiarity Matrix, we offer this diagram (Figure 6).

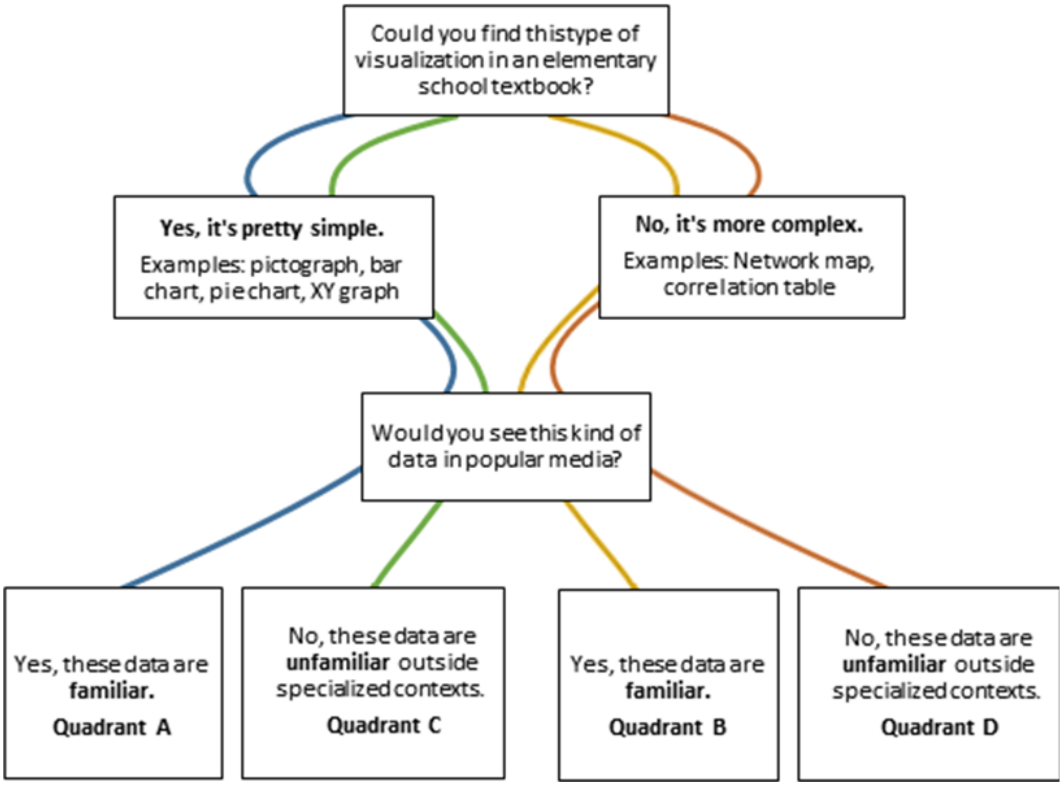

Figure 6. Simplicity Familiarity Matrix. [Color figure can be viewed at wileyonlinelibrary.com]

First, we suggest determining if the visualization is simple or complex. Would you find the visualization in an elementary school text book? (Examples that would fit this context might include a pictograph, bar chart, or XY graph.) Further, does it use a common reference system or simply calculated dichotomous data? If it involves more features (e.g., 3 or more axes or nested bars) or is not something the average person might encounter in their daily activities, it would be considered complex. After this is determined, we suggest determining the complexity of the data. Is it found in the popular media or is it limited to specialized fields?

Visualizations that fall in Quadrant A can be used throughout an exhibition with minimal scaffolding. If the data is familiar, curators may consider more complex data visualizations, as audiences fill face fewer conceptual barriers. Conversely, less familiar data could be used if the visualization is simple. Depending on the level of familiarity with the data and visualization, visitors may need hints or significant contextual framing in order to fully understand the conceptual meaning of less-familiar data or complex data visualizations.

Recommendations

To aid in visitor understanding of data visualizations, the findings from the *Sense Making of Big Data* project suggest that interpretation in informal learning settings may be best supported by (1) employing well-designed, easy-to-understand data visualizations, (2) providing appropriate support and guidance, (3) offering multiple modes of access, and (4) making data visualizations relevant to visitors, e.g., via personalization.

While these findings echo established principles of strong interpretation practice, they can serve as a reminder for the seasoned professional and an introduction for emerging professionals and students in the field. Below, we offer data that underpin each principle, along with specific recommendations for supporting learners with visualizations from across the Simplicity-Familiarity Matrix.

Design and/or Use Data Visualizations That are Clear and Easy to Understand

For any given interpretive project, professionals are immersed in the data and graphics, and they can develop assumptions that museum visitors are able to make sense of complex visualizations; these assumptions may or may not be true. To ensure visitor comprehension, focus on the fundamentals of data visualization; highlight essential elements of data visualizations including a clear title, legend, labels and defining key terms. To encourage further study, information on data and algorithms used together with source credits and even short bios of authors might be provided. This is particularly important for data visualizations categorized as complex (i.e., those that fall in Quadrants B and D). Consider highlighting the XY axis in data visualizations, as this is a basic element in many different types of data visualizations. Additionally, museums may wish to consider undertaking front-end studies that include mock-ups of proposed data visualizations and have visitors talk out loud as they make meaning from the visualization in order to ascertain if the majority of visitors understand the data visualizations included in an exhibit.

Museum visitors bring their prior knowledge to each visualization. In the *Gapminder 2012* data visualization, there was some confusion over the term “country.” When the second layer was added, more than half of the individual/groups recognized that the bigger circles added at this layer represented countries with larger populations; however, there were individuals who interpreted the word “country” as rural. Additionally, both adults and children were confused by the term “axis,” which was used to describe the position of countries on the graph. A child remarked that “the earth spins on its axis.” A legend defining the symbols and labels used, could benefit all visitors.

Even with simpler visualizations, it is supportive to condense information where possible; ensuring that every element of the data visualization adds to visitor understanding, not overload. In the face of elements that do not add critical meaning, visitors can be overwhelmed by the amount of information presented on a data visualization and may have trouble determining what is most important. For example, the *Competitive Eating Records* graphic included several pictures of food. While this inclusion received some positive comments from participants, such as that “pictures provide context” and that images mean that the learner “can visualize,” a similar number of adults felt the pictures did not add to the visualization – including one observation that the pictures were a source of confusion that distracted learners from fully comprehending unfamiliar data.

Provide Visitors Support, Give Guidance About How to Read Visuals, and Allow for a Visitor’s Flexible Time Investment

Given the range of familiarity with data visualizations observed across the five sites, visualizations that present multiple levels of complexity are most likely to meet the interpretive needs of both the novice and the knowledgeable visitor. In practical, constructionist terms, this would mean beginning at a basic level, with the potential to increase in complexity in logical, sequential steps to the extent that the visitor is interested.

To move visitors progressively toward stronger understanding of data visualizations, and greater independence in reading and understanding of a variety of data visualizations, museum professionals might therefore consider supporting selected data visualizations in an exhibit by unpacking the layers and identifying key elements. In this study, visitors constructed the data visualizations one layer at a time and were more likely to use cumulative reasoning, making deeper meaning. This suggests that building from simpler to more complex data visualizations can increase guest understanding of complex data visualizations, and for an exhibit where a key message is supported by complex data, these findings can provide a way of integrating those data into a meaningful presentation.

Additionally, providing the opportunity to construct a data visualization allows visitors to spend as much (or as little) time with a given exhibit as is appropriate to their experience goals and learning needs. These modular experiences can both stand alone in short segments or experienced in sequence for those willing to invest the time. With so many options in a typical informal learning setting, plus the added challenge of navigating the space together with their group, a visitor's time spent with any one exhibit can be limited. Respecting a visitor's time while providing several paths to understanding, may enable visitors to make sense of a data visualization. Even if a visitor does not completely construct a data visualization, they have increased or reexamined their previous understanding.

Accommodate a Variety of Learning Preferences

Museums can enhance learning for those who prefer to learn through a range of different sensory modes (e.g., visual, auditory, or kinesthetic) by presenting data visualizations in a variety of static and dynamic formats, which allow visitors to manipulate the data or add their personal data to the data set. A prototype exhibit for the Big Data project, created and facilitated by the Science Museum of Minnesota, encouraged visitors to participate in the creation of the data visualization by adding a sticker to the appropriate location on a map, indicating where they lived, found visitors eager to participate in the creation of a data visualization. In addition to the typical static data visualization mounted and displayed, dynamic displays can include on demand labels, definitions, and links to additional information, or more complex graphics. Other strategies for supporting learners could include in-person facilitators: museum educators could offer real-time support by explaining the different elements of the visualization (i.e., the reasoning behind a particular number scale, the importance of color in a heat map, or the layers of a layered bar graph), and offering responsive feedback as the visitor constructs their data visualization.

Connect Data Visualizations to Visitors' Daily Lives and Personal Experiences

In the *Sense-Making of Big Data* construction study, the most misunderstood visualization was *Padgett's Florentine Families*, possibly because it referenced Italian history, which was not a common frame of reference for the majority of these museum visitors. This finding highlights the need to connect data visualizations to visitors' daily lives.

Connecting data visualizations to the visitor can be done through words or pictures. In this study, the majority of participants recognized that the *Competitive Eating Records* graphic referenced food, possibly because of the familiar food names and/or the iconic food graphics displayed on the graph. Visitors understood the basis of the *Gapminder 2012* visualization based on terms familiar to most, such as "life expectancy" and "income per person."

The *US Airport Traffic and Chicago O'Hare Connecting Flights* visualization may have been the easiest to construct meaning due to the location of the data collection. Much of the data collection for this study took place in the Midwest, so Chicago was a familiar referent. Additionally, many, if not all, of the adult participants had flown, and some cited spending time in O'Hare Airport on layovers. Children who did not have this frame of reference, because they are less travelled or unaware of the names of Chicago's airports, were helped by the adult with which they were visiting. Adults would offer potential frames of reference for the child, such as "remember when we went to X, this is where we stopped to change planes," or "you know how you like airplanes. Well O'Hare is one of the places where lots of planes fly to and from."

Other studies have found that visitors find value in participating in the creation of data visualizations (Wojton and Heimlich 2015). Touch screen displays that encourage visitors to participate in brief "research" studies, logging their observations when prompted, and creating a simple graphic of their data vs. a larger set of all visitors' data allow individual learners to have a personal experience with data collection and visualization. This type of engagement can also be facilitated using cost-effective methods and tools. For example, attendees at the Aquarium of the Pacific's NASA Night event were invited to help facilitators plot the location of kelp forest animals in the Aquarium's *Blue Cavern* exhibit on a nearby whiteboard and to plot the relative popularity of favorite animals at the Aquarium by coloring squares on a large sheet of butcher paper.

CONCLUSION

This study found visitors were unfamiliar with very specialized data and that data visualizations with more than average complexity were difficult for visitors to interpret. Despite this, the data also suggested that museum practitioners may assume a basic level of familiarity with certain types of data, as well as a common understanding of simple data visualizations. To ensure all visitors are able to understand the visualizations within exhibits and participate in programs, practitioners must provide basic information with opportunities for those more knowledgeable to move beyond the basic to something more complex. In order to ensure all (novice and advanced) visitors understand the data visualizations used throughout the museum, scaffolded experiences that provide opportunities to understand unfamiliar data or complex data visualizations are essential. Use of the Simplicity-Familiarity Matrix can assist practitioners in determining the level of scaffolding necessary to ensure visitor understanding. In addition, prototyping experiences that include data visualizations can allow curators an opportunity to understand how visitors interpret the data before text and exhibits are finalized.

The ubiquity of big data suggests that visitors to museums and science centers will continue to encounter complex data visualizations in their daily lives, and as informal learning professionals ourselves, we encourage the inclusion of data visualization in interpretation that seeks to help people make sense of the world around them. While future research can help us understand which elements of data visualizations are *most* important to construction of meaning, the findings described above illustrate that a few key principles support learners in navigating the interplay of these elements – for both big data and simpler data sets. While this study found evidence that indicates visitors increased their understanding of the data while co-creating the visualization, additional study could deepen this understanding.

We expect that the general practice of building interpretive meaning from basic frameworks up (to more complex variables and/or unfamiliar subjects) will help learners of all ages and skill levels navigate the visual information they encounter, particularly in informal learning settings. For practitioners, this discussion is intended to provide some effective starting points for decision-making about what types of data visualizations and which data sets are included in museum interpretation. Further, we anticipate that the principles described above will support interpretation professionals as they work to help public audiences make sense of data visualizations. Ultimately, interpretation of large data sets concerns not just the presentation of data, but the presentation of many layers of expression; by starting with basics and revealing those layers, rather than obscuring them, we invite our audiences not just to make their own meaning, but to join us in a process of building it. **END**

ACKNOWLEDGMENTS

This research was completed with support from the National Science Foundation #1223698 and NASA #NNX12AL16G

REFERENCES

Alper, B., N. H. Riche, F. Chevalier, J. Boy, and M. Sezgin. 2017. "Visualization Literacy at Elementary School." In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 5485–97. New York: Association of Computing Machinery. Accessed February 5, 2018. Retrieved from https://doi.org/10.1145/3025453.3025877.

Azzam, T., S. Evergreen, A. A. Germuth, and S. J. Kistler. 2013. "Data visualization and evaluation." *New Directions for Evaluation* 139(1): 7–32.

Börner, K. 2015. *Atlas of Knowledge: Anyone Can Map*. Cambridge, MA: The MIT Press.

Börner, K., A. V. Maltese, R. N. Balliet, J. E. Heimlich, and S. Uzzo. 2015. "Data Visualization Literacy of Youth and Adult Science Museum Visitors." Session presented at the American Education Research Association Conference, Chicago, IL, April 16–20, 2015.

Börner, K., and T. Polley. 2014. *Visual Insights: A Practical Guide to Making Sense of Data*. Cambridge MA: The MIT Press.

Börner, K., A. Maltese, R. N. Balliet, and J. Heimlich. 2016. "Investigating Aspects of Data Visualization Literacy Using 20 Information Visualizations and 273 Science Museum Visitors." *Information Visualization* 15(3): 198–213.

Bright, G. W., and S. N. Friel. 1998. "Graphical Representations: Helping Students Interpret Data." In *Reflections on Statistics: Learning, Teaching, and Assessment in Grades K-12*, edited by S. Lajoie., 63–88. New York: Routledge.

Canning, K. 2013. "Put a Frame on It: Contextualizing Climate Change for Museum Visitors." Unpublished master's thesis, University of Washington. Accessed September 25, 2013. Retrieved from http://hdl.handle.net/1773/23527

Doering, Z. D., and A. J. Pekarik. 1996. "Questioning the Entrance Narrative." *Journal of Museum Education* 21(3): 20–3.

Hachet, M., J.-B. de la Riviere, J. Laviole, A. Cohe, and S. Cursan. 2013. "Touch-Based Interfaces for Interacting with 3D Content in Public Exhibitions." *IEEE Computer Graphics and Applications* 33(2): 80–5.

Hein, G. E. 1998. *Learning in the Museum*. New York: Routledge.

Hinrichs, U., H. Schmidt, and S. Carpendale. 2008a. "EMDialog: Bringing Information

Visualization into the Museum." *IEEE Transactions on Visualization and Computer Graphics* 14(6): 1181–8.

Hinrichs, U., H. Schmidt, T. Isenberg, M. S. Hancock, and S. Carpendale. 2008b. "BubbleType : Enabling Text Entry within a Walk-up Tabletop Installation." University of Calgary. Accessed October 30, 2013. Retrieved from https://research-repository.st-andrews.ac.uk/handle/10023/7414.

Hohenstein, J., and T. Moussouri. 2018. *Museum Learning: Theory and Research as Tools for Enhancing Practice*. New York: Routledge.

Kafai, Y. B., and M. Resnick (eds.). 1996. *Constructionism in practice: Designing, thinking, and learning in a digital world*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kaminski, J. A., and V. M. Sloutsky. 2013. "Extraneous Perceptual Information Interferes with Children's Acquisition of Mathematical Knowledge." *Journal of Educational Psychology* 105(2): 351–63.

Kim, B. 2001. "Social constructivism." In *Emerging perspectives in learning, teaching and technology*, edited by M. Orley, 55–61. Zurich, Switzerland: The Global Text Project.

Konold, C., and T. L. Higgins. 2003. "Reasoning about Data." In *A Research Companion to Principles and Standards for School Mathematics*, edited by J. Kilpatrick, W. G. Martin and D. Schiffer, 193–215. Reston, VA: National Council of Teachers of Mathematics.

Korn, R. 2015. *Formative evaluation round one: New York at its core*. Technical report. Alexandria, VA: RK&A, Inc. Accessed January 26, 2018. Retrieved from http://informalscience.org/sites/default/files/2015_RKA_MCNY_NY%20At%20Its%20Core_Formative%20Evaluation_final_dist.pdf.

Kostakos, V., H. Kukka, J. Goncalves, N. Tselios, and T. Ojala. 2013. "Multipurpose Public Displays: How Shortcut Menus Affect Usage." *IEEE Computer Graphics and Applications* 33(2): 56–63.

Krumhansl, R., C. Peach, J. Foster, A. Busey, I. Baker, and J. DeLisi. 2012. *Visualizing Oceans of Data: Educational Interface Design*. Waltham, MA: EDC.

Lincoln, Y. S., and E. G. Guba. 2013. *The Constructivist Credo. Igarss 2014*. Walnut Creek, CA: Left Coast Press.

Ma, J., I. Liao, K.-L. Ma, and J. Frazier. 2012. "Living Liquid: Design and Evaluation of an Exploratory Visualization Tool for Museum Visitors." *IEEE Transactions on Visualization and Computer Graphics* 18(12): 2799–808.

Papert, S., and I. Harel. 1991. "Situating constructionism." In *Constructionism*, edited by S. Papert and I. Harel., 1–11. Norwood, NJ: Ablex.

Roth, W.-M., and G. M. Bowen. 1994. "Mathematization of experience in a grade 8 open-inquiry environment: An introduction to the representational practices of science." *Journal of Research in Science Teaching* 31(3): 293–318.

Thudt, A., U. Hinrichs, and S. Carpendale. 2012. "The Bohemian Bookshelf: Supporting Serendipitous Book Discoveries through Information Visualization." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1461–70. New York: Association of Computing Machinery. Accessed October 30, 2013. Retrieved from https://doi.org/10.1145/2207676.2208607.

Tranby, Z., M. A. Wojton, and J. Heimlich. 2013. *Visitor Understanding of the Term Big Data*. Minneapolis: Science Museum of Minnesota.

Wojton, M. A., and J. Heimlich. 2015. *iSaveSpecies—Great Apes Summative Evaluation Report for Zoo Atlanta*. Technical report. Columbus, OH: Lifelong Learning Group. Accessed April 1, 2015. Retrieved from http://www.informalscience.org/sites/default/files/Atlanta_iSaveSpecies–Great%20Apes%20Summative_Final.pdf

Wojton, M. A., S. Palmquist, V. Yocco, and J. Heimlich. 2014. *Meaning making through data representation: Construction and deconstruction*. Technical report. Columbus, OH: Lifelong Learning Group. Accessed November 10, 2014. Retrieved from http://informalscience.org/sites/default/files/2014-11-13_Big_Data_Construction_Deconstruction_Report.pdf

Wolfe, P. J. 2013. "Making Sense of Big Data." *Proceedings of the National Academy of Sciences of the United States of America* 110(45): 18031–2.