
How Prior Knowledge Affects the Processing of Visualized Data

Jessica Hullman

University of Michigan
School of Information
105 S. State St.
Ann Arbor, MI 48109 USA
jhullman@umich.edu

Abstract

Research on how to align information visualization interfaces with user characteristics has oft focused on general visual perceptual models starting from low-level feature processing. Yet cognitively-driven “top-down” factors also carry effects. I describe how one such factor—prior knowledge—affects visualization perception and interpretation. This includes highlights from my own research on two knowledge influences: impacts from social information and temporal dependency in judgments. Noting interacting variables enables accounting for knowledge in interface design.

Author Keywords

Authors’ choice; of terms; separated; by semi-colons

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction

Models of graph comprehension in information visualization have traditionally been concerned with how comprehension arises from “bottom-up” perception: prioritization of visual features, color vision, and perceptual faculties for different types of visual judgments, to name a few (e.g., [14]). Yet various studies suggest that “top-down” effects based on prior knowledge and beliefs are just as important, influencing the visual features that are encoded as well as the interpretations formed (see Figure 1) [10]. Consider re-examining a scatterplot to judge a correlation between variables *after* computing the statistical association: it becomes difficult to perceive the plot independently of the statistical information. Other examples include a common belief that graphs tend to convey “scientific” information that can consequently magnify the salience of visualized information in decision-making, or the way that domain knowledge can guide what information in a visualization is brought to the user’s focus.

While prior knowledge effects on graph comprehension have been demonstrated in educational settings (e.g., [4]), the increased prevalence of visualizations in online

Knowledge Relevant to Visualization Interfaces

Beliefs about graphs:

"Graphs are scientific" [13].

Beliefs about graph types:

"Line graphs show trends" [15].

Knowledge of the topic:

"I've seen that variable before" [11].

Beliefs from culture: "A positive (or negative) trend is most likely to reverse in direction" [8].

Beliefs from social information:

"They saw X, and so do I" [6].

Knowledge from prior views:

"The trend in this visualization looks a lot like the last one" [7].

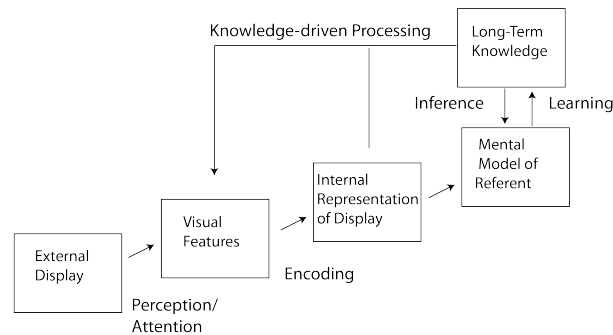


Figure 1: Model of display comprehension proposed by [10].

environments from news to social media make it important to consider how knowledge influences the interpretation of visualizations by diverse non-domain experts. Existing work has distinguished between prior knowledge effects related to graphical formats [15] and related to content (e.g., knowledge that boys are typically taller than girls [11]). I integrate factors explored in my own research: effects of social information and information gained earlier in an interactive visual analytics session. Motivated by the potential for diverse online users to produce valuable insights into visualized data, I discuss how knowledge effects and their interactions with other user and task features affect interpretation quality.

Effects of Prior Knowledge About Format

Graphical Formats

Expectations associated with specific graph types (such as how a given format tends to be used) can focus user attention on particular patterns or comparisons. Zacks and Tversky [15] found that line graphs tend to be interpreted as depicting trends, even when the data presented is categorical. For example, when viewing a graph of boys' and girls' heights, participants made statements such as "as people become more male...", even though this interpretation is clearly flawed. If multiple lines are shown in the same graph, individuals

will typically focus on comparisons between relative slopes of those lines (i.e., one line is increasing, another is decreasing), and pay less attention to the relative positions of those lines. In contrast, when individuals view bar graphs, they tend to compare the relative difference between bars that are grouped together [11]. Such effects are likely to be associated with the graphical and statistical literacy of users. Online visualization systems that promote visualization creation by general web users such as IBM's Many Eyes, for example, may be at risk from users deviating from familiar graph conventions in creating visualizations (e.g., depicting trends with bar charts). Features for inferring whether an x-variable is temporal, for example, could help prevent mismatches.

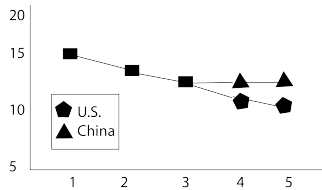
A more general type of format expectation concerns the way visualizations are often associated with scientific analyses. Chua et al. [1] asked participants to state their willingness-to-pay for products (toothpastes or tires) that would reduce risks (tooth decay or tire blowouts, respectively). Graphical presentations led subjects to value the reductions to risk more than did text and to report greater affective responses to visualized "risk". This effect appears to be caused by several features of graphs that may interact to emphasize behavioral effects: 1) visual differences being perceived as larger, and 2) the way graphs often cue a viewer that the data are scientific [13]. An implication of these results is that systems should reserve visualization for more reliable information.

Effects of Prior Knowledge About Content

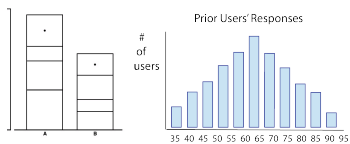
Domain Knowledge

Many general domain knowledge effects apply to visualization. Individuals tend to be more critical of

Examples



A. Cultural Norms: Chinese and American users predict future points of 3 point trend [8].



B. Social Information: Input from prior users (right) biases visual judgments of bar graphs (left) [6].



C. Prior views: Most recently-viewed visualization affects judgment for entire series [7].

information that is incongruent with their beliefs [11], and this relevant to visual analyses of 2D scatterplots. I am currently researching whether particular graphical formats can better overcome such biases. Familiarity with the content of the information being depicted can influence comprehension in more subtle ways as well, guiding eye movements to the correct patterns [5] and stimulating the development of mental models based on perceived gaps in knowledge [4]. Specific to visualizations, Shah & Freedman [11] find that when data depicted information familiar to viewers, they were better able to draw appropriate inferences from graphs, rather than simply focusing on salient information as did non-familiar users. Hegarty et al. [3] provide further support, examining how bottom-up and top-down processes interact in using complex weather map interfaces. They find that visual salience affected accuracy only when prior task-relevant knowledge was in place, with eye fixations being fairly unrelated to the display design. Hence, good visualization design facilitates performance not just by guiding where viewers look but also by supporting processing of the task-relevant visual features at a given location.

Cultural Norms

While implicit or unconscious influences of culture are known to affect visual attention, more conscious expectations that vary with culture can also influence how visualized data trends are perceived. Ji et al. [8] presented line graphs indicating trends (e.g., the death rate for cancer), and asked Chinese and Americans to predict the probability that the trend would go up, down, or remain the same in comparison to the last point on the line, or for a prediction of the next two points in the series (sidebar A). For patterns with

changing slopes, American were less likely to predict change in the rate of change, but when things are changing in a particular direction, they are more likely to predict change in direction of the prior movement.

Social Information

In prior work, I found that summarizing others' perceptions in a basic graph task (e.g., proportion judgment) affects subsequent users' visual judgments [6] (sidebar B). Aggregate responses to basic visual judgments thus depend on the quality of social information. If initial information is biased, such as when systematic bias affects a visual task, then the biased judgment can propagate across the group. Designing visualization systems that withhold social information until a reliable signal is available, or introduce variation into the visual data to be judged between users (e.g., data bootstrapping) may help.

Prior Interactions with Similar Data

Decision science has shown how prior knowledge (in the form of contextual dependencies like the order of presentation, or the presence of an irrelevant prime) affects processing of subsequent information [9]. My colleagues and I used online experiments to investigate whether such effects also apply to visualization processing [7]. For instance, does the order in which visualizations are viewed sequentially (such as in an interactive visualization setting with multiple views) affect overall judgments for the series? (See sidebar C). We observed strong order effects that suggest a recency bias where the visualization seen directly prior unduly affects an overall conclusion. Further tests are necessary to disambiguate, but it is plausible that a user recalls a visualization seen earlier in the series as

Interacting Variables

Statistical literacy:

Awareness of the correct way to interpret depicted comparisons (such as the concept of statistical significance).

Graphical literacy:

Awareness of the correct way to interpret a design, such as the way that the two boxes of a boxplot contain the same number of observations.

Saliency of design: Degree to which organization of visual information makes relevant patterns easily visible.

Interactive features: When prior domain-relevant knowledge is not available, animation can help users infer what to look for [4]; yet too many interaction choices may lead novices astray.

Text Framing: Users' responses to given prompts or questions can differ depending on expectations cued by the text phrasing.

more similar to the current visualization than it actually is. Further, an interaction was observed between the order of information and how a question is framed. We found that when we rephrased a question that guided the user's analysis of the visualization (from "Do the deaths appear to result from a single source?" to "Do the deaths appear to result from multiple sources?"), responses to the visualizations significantly differed. Here, effects may stem from prior knowledge in the form of implicit beliefs that the visualized evidence will be more likely to represent positive (confirmatory) evidence for the hypothesis suggested by the question rather than negative. This may cause users to selectively search for confirmatory evidence (see confirmation bias literature).

Conclusion

Prior knowledge related to the format or content of a visualization interface can significantly affect perception and comprehension. By acknowledging likely effects and their interactions, we can enable deeper understanding of how to design effective visualization interfaces for diverse online audiences.

References

- [1] Chua, H. F., Yates, J. F., & Shah, P. Risk avoidance: Pictures versus numbers. *Memory & Cognition* 34, 2, (2006), 399-410.
- [2] Cleveland, W.S., & McGill, R. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *J. of the American Statistical Association*, 79, 387, (1984), 531-554.
- [3] Hegarty, M., Canham, M., & Fabrikant, S. Thinking about the Weather: How Display Saliency and Knowledge Affect Performance in a Graphic Inference Task. *J. of Experimental Psych.* 36, 1 (2010), 37-53.

[4] Hegarty, M., & Kriz, S. Effects of Knowledge and Spatial Ability on Learning from Animation. In *Learning from Animation*, Lowe, R, & Schotz, W. (2007), 3-29.

[5] Henderson, J., & Ferreira, F. *The Interface of Language, Vision, and Action: Eye Movements and the Visual World*, Psychology Press, 2004.

[6] Hullman, J., Adar, E., & Shah, P. The Impact of Social Information on Visual Judgments. *ACM CHI* (2012), 1461-1470.

[7] Hullman, J., Krupka, E., & Adar, E. A Bootstrap Mechanism for Crowdsourcing Visual Analysis. Long article in preparation.

[8] Ji, L., Nisbett, R., & Su, Y. Culture, Change, and Prediction. *Psych. Science* 12, 6 (2001), 450-456.

[9] Kahneman, D., Slovic, P., and Tversky, A., eds. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge University Press, 1982.

[10] Kriz, S., & Hegarty, M. Top-down and bottom-up influences on learning from animations. *International J. of Human-Computer Studies* 65, (2007), 911-930.

[11] Sanitioso, R., & Kunda, Z. Ducking the collection of costly evidence: Motivated use of statistical heuristics. *J. of Behavioral Decision Making*, 4, (1991), 161-178.

[12] Shah, P., & Freedman, E. Bar and Line Graph Comprehension: An Interaction of Top-down and Bottom-up Processes. *Topics in Cog. Sci.* 3, 3 (2011), 560-578.

[13] Smith, L. D., Best, L. A., Stubbs, D. A., Johnston, J., & Archibald, A. B. Scientific graphs and the hierarchy of the sciences: A Latourian survey of inscription practices. *Social Studies of Science* 30, (2000), 73-94.

[14] Ware, C. *Visual Thinking: for Design*, First Edition. Morgan Kaufmann, 2008.

[15] Zacks, J., & Tversky, B. Bars and lines: A study of graphic communication. *Memory and Cognition* 27 (1999), 1073-1079.