Memorability of Enhanced Informational Graphics: the effects of design relevance and chart type on recall

Alyssa Peña

Research Assistant
Department of Visualization
Texas A&M University

mupena17@tamu.edu

Eric D. Ragan*

Assistant Professor
Department of Computer & Information Sc. & Engineering
University of Florida

eragan@ufl.edu

Lane Harrison

Assistant Professor
Department of Computer Science
Worcester Polytechnic Institute

ltharrison@wpi.edu

**corresponding author*

Abstract /

Design enhancements are often added to charts, signage, and infographics to help garner attention or communicate a message. Despite concerns of design enhancements distracting from the underlying information, prior studies have contributed evidence that enhancements and even simple decorations can improve memory of visual displays. However, there is limited empirical knowledge about how the type of aesthetic enhancements influences memory, and what informational and data elements are remembered. We conducted an experiment testing chart types (line, pie, and bar), presence of color, and whether added enhancements were contextually related to the chart's data topic. The findings show relevant enhancements did improve recall of title and thematic elements, but enhancements did not significantly affect recall of specific data values. This suggests relevant enhancements can improve memorability of some chart content, but only if design styles are chosen well to match the information. Recall of chart topics for unrelated embellishments was worse than non-enhanced charts, which suggests that enhancements can distract or interfere with memorability if the viewer does not understand a meaningful connection between informational topic and design modifications.

Keywords /

visual design; information visualization; memory and recall; infographics; empirical study

INTRODUCTION

Infographics have proven to be highly effective means of communicating data in meaningful and understandable ways. Various forms of signage (e.g., posters, electronic messages, billboards) commonly take advantage of carefully crafted infographic designs to increase appeal and better inform or persuade viewers. Infographic signage include displays designed to educate the public about safety recommendations, civic issues, or useful information for travel planning. They are also used in marketing signage for appealing presentations of information related to products or services. Electronic signs can take further advantage of infographic displays that regularly update data to provide individualized or timely information to viewers.

While there has been interest in public use and interpretation of visualization with aesthetic enhancement and manipulations (Claes and Vande Moere, 2013; Holmes, 1984; Skog et al., 2003), visualization experts have long warned about the risks of "embellishment" or "chart junk" detracting from a visualization's ability to clearly communicate information (see Tufte, 1983). Despite the potential downsides to added stylistic enhancements, multiple studies have contributed evidence that visual enhancements can improve memory of visualizations (e.g., Bateman et al., 2010; Borkin et al., 2013; Quispel et al., 2016). One interpretation of such results is that prudent enhancements can make visuals more effective at communicating a message or making information more memorable.

Studies have found evidence that embellishment can influence the memorability of visual aspects of information presentations, but the extent to which embellishment can affect memorability of the information itself is less certain (Few, 2011). While several studies have investigated the relationship among the types of aesthetic embellishments and what elements of the visualization are remembered (Borgo et al., 2012; Borkin et al., 2013; Haroz et al., 2015), the design space is expansive, the result of the variance in types of fundamental data representations and embellishment strategies.

As such, limited empirical knowledge regarding what type of information is recalled from different visual properties of infographics exists and a clearer understanding of the implications of design enhancement for visualization memorability requires further study. In this paper, we present a controlled experiment to evaluate how added enhancements affect the memorability of a visualization's underlying data and informational elements. The user study follows a two-part design to evaluate recall of multiple types of chart information after one week. In addition to comparing the presence and absence of embellishments for multiple chart types, the study evaluates whether memorability is affected by the relevance of the enhancements to the data topic.

RELATED WORK

The widespread use of infographics and other visualizations in information environments has led to multiple lines of inquiry in the visualization community, including interview studies and quantitative experiments. These prior works include several quantitative studies on embellishments in charts, which inform several dimensions of our current study. Because the study presented here focuses on design effects for enhanced information displays, we prioritize background research in information visualization. A common goal for information visualization is to present data in such a way that it is readable, accurate, and understandable (Kosara, 2007). Graphical perception studies aligning with this goal have studied how different forms of fundamental visual encodings, such as mark position, length, size, and shading, influence how accurately people can interpret data values (Cleveland and McGill, 1984; Heer and Bostock, 2010; Saket et al., 2018; Szafir, 2017). Beyond graphical marks, graphical perception studies have also explored peoples' ability to extract statistical measures from visualizations (Beecham et al., 2017; Correll and Heer, 2017b; Harrison et al., 2014; Rensink and Baldridge, 2010; Wickham et al., 2010). Despite the growing trend of infographic-style visualizations in areas of daily life such as news and social media, comparatively few studies have targeted the use of infographics to inform, educate, or persuade casual observers (Kellaris and Machleit, 2016).

Design Enhancement, Infographics, and Memorability

Prior research has also focused on defining and measuring higher-level concepts in visualization; examples include creative data representations, visual enhancements, and their impact on users' performance and behavior (Bateman et al., 2010; Borkin et al., 2013; Bylinskii et al., 2017; Haroz et al., 2015; Lahrache et al., 2018; Ruchikachorn and Mueller, 2015; Zacks et al., 1998). We draw on the results and methodologies of several of these studies as the basis for the main questions explored in this paper, namely whether particular styles of design enhancements might influence a persons' ability to remember details of a chart. For example, Batemen et al. (2010) compared enhanced and plain charts by testing the possible influence of embellishments on their interpretation, accuracy, and long-term recall. In a study with two sessions, participants first looked at charts and answered a series of questions. They were then instructed to leave for a predetermined amount of time before returning and being asked similar questions about the charts they had studied; these questions focused on common facets of the chart, including its subject and data values, trends, and changes. After participants left, they were called back either five minutes or two to three weeks later for a recall session and when they returned, participants were again asked to describe trends and messages in the charts they had seen. The results of the study showed that while accuracy of data interpretation was similar between plain and enhanced charts, enhanced charts led to better long-term recall scores.

Other studies have found relationships between chart styles and memorability. In a study by Borkin et al. (2013), participants viewed a series of charts and were asked to specify when they saw repeated charts. The results highlighted how people were more likely to recall colorful visualizations, those using unique visual representations of data, and those using design enhancements that could possibly interfere with the accuracy of interpretation. In a later study, Borkin et al. (2016) revisited memorability by breaking down structural components of infographics further, with results suggesting that clear titles, supporting text, and redundancy help with the recall of visual representations of data they had seen previously.

Highly relevant to our research, Borgo et al. (2012) studied memorability of data charts with enhancements. They concluded that additional enhancements can increase memorability of chart information, but participants also took longer to interpret the charts prior to recall—meaning they had longer exposure. We observe that most prior related studies have focused on infographics gathered in the "wild," from existing visualizations in media and on the internet. In contrast, we adopt a study methodology with controlled variation of embellishments in charts to allow for a more systematic examination of the possible benefits and risks of using enhancements in charts. Haroz et al. (2015) is perhaps most applicable to the investigation of enhancement's relevance through controlled experimentation, as the research studied various approaches to incorporate icons and representative imagery into bar charts. Variations of embellishments included the use of a stretched icon behind bar charts, having the bars made of a row of small icons, or having a single large icon in place of each bar. This study considered the relevance of enhancements by including icons that either exactly matched an item name (e.g., a dog icon for a dog category) or was related to the category but did not match an item (e.g., a fish icon for a dog category). In this way, the mismatched icons would demonstrate a difference while still preserving relevance to a higher-level category (e.g., animals). The study found greater errors for recalling specific values with the presence of mismatched icons. In addition, their research found evidence that compared to simple charts without embellishment, participants gave more attention to bar charts with pictographic enhancements added.

The prior research on memorability of visualizations and infographics provides a strong foundation for continued research due to the number of observed effects of enhancements, but further evaluation is needed to develop a fuller understanding of the implications. Our research investigates goals similar to Borgo et al. (2012) and Haroz et al. (2015) with several key differences. First, our experiment includes relevance of enhancements as a controlled experimental factor. Haroz et al. considered relevance via icons that either (a) exactly matched bar chart categories or were (b) the wrong icons but were still related to the general topic of the chart. The study found the relevant-but-wrong visual embellishments were confusing and had negative effects on memorability, but the preservation of the relevance to the topic may have had a distraction or interference effect. Our study, therefore, aims to investigate entirely unrelated embellishments. Further, we study enhancements other than strict iconographic encodings with discrete values (e.g., five icons represent the value of five). In addition, our study provides a controlled comparison of fundamental chart types (bar, line, and pie charts). While Borgo et al. (2012) considered a broad variety of visual designs for breadth, chart design was not controlled as a factor in the study, whereas we seek to explicitly test for effects of different visual representations.

Our study also evaluates recall after a longer period of time (one week compared to seconds after chart viewing). Testing after an extended time has benefits for relevance in many realistic contexts (remembering data/charts days after reviewing them rather than immediately after), which may be the reason why some prior researchers have opted for longer term studies (see Bateman et al., 2010). Memory research has shown evidence of significant drops in recall ability over a period of days (Atkinson and Shiffrin, 1968). Moreover, our study design is influenced by memory research that has shown that relevance or meaningfulness of information influences how easily it can be recalled (Chase and Simon, 1973; Craik and Tulving, 1975; Smith and Graesser, 1981). This is especially important for our study of embellishment relevance.

Furthermore, we included data inspection tasks to encourage the acquisition of embedded information. Because the application of information is thought to reinforce memory, a longer duration is more important for a meaningful evaluation of differences (Nuthall, 2000; Ritchie and Karge, 1996). With the importance of systematic empirical research and separate independent experimentation, the presented research aims to further assess the topic of memorability while expanding the body of knowledge with increased attention to differences in chart types, differences in embellishment relevance, and prolonged time before recall assessment. Taken together, prior works at the intersection of visualization perception, design enhancement, and facets such as memorability and aesthetics yield a series of open challenges and methodologies and framings that inform our present study.

METHOD

We conducted a controlled experiment to evaluate how visual design enhancements and chart properties affect the memory of informational elements of infographics.

Research Goals and Hypotheses

Rather than focus solely on the memory of a visual message, our goal was to investigate whether memory of a chart's informational elements is influenced by chart design and aesthetic enhancements. In this study, chart information includes data values, chart titles, and overall topics or trends. By considering the effects of chart properties, we study whether viewers are internalizing the presented information or the imagery itself. Our research also evaluates effects due to the contextual nature of enhancement, focusing on visual additions that add supplemental imagery without manipulating the fundamental chart design. For this study, we use the term embellishment to emphasize that the design enhancements studied in this project are simple visual additions that do not modify the underlying data representation. In other words, the study focuses on the addition of "decorative" elements rather than cover the broader scope of design enhancements or alterations possible through graphic design.

We designed the experiment to assess whether the addition of embellishments helps improve memorability, or if it makes a difference if enhancements are contextually relevant to the chart's data and topic. Further, we were also interested in studying whether basic chart types and properties influence memorability of informational elements. More specifically, we wanted to test different fundamental chart types (e.g., line charts, bar charts, pie charts), because certain visualization formats may be better suited for certain data types and tasks and different shapes of data representations can influence interpretation (Jardine et al., 2019). The experiment accounts for possible interactions with color, since previous research has found evidence that chart color can affect memorability, and contribute semantic information for data categories (Borkin et al., 2013; Lin et al., 2013).

The following hypotheses were based on our goals and prior work:

H1: Contextually relevant enhancements will increase memorability of chart information compared to unrelated embellishments.

H2: Fundamental chart types (line, bar, pie) might influence memorability.

H3: Color images will be more memorable than black-and-white data graphics.

Experimental Design

The experiment followed a 3x3x2 mixed design where three chart types and three embellishment types were varied within subjects, and two levels of color were varied between subjects. For chart types, we tested line, pie, and bar charts in order to provide variety of different fundamental visualization designs for data representation. Figure 1 shows examples of the three chart types. Each row consists of three versions of the same chart modified according to the three different embellishment types.

The three types of embellishment (related, unrelated, and plain) controlled the style and decoration to the chart. Charts with related embellishments were designed to include supplemental visual elements that were thematically related to the topic of the chart. The left-most column of Figure 1 provides examples. In contrast, unrelated embellishments included stylistic modifications that were not related to the chart information (see middle column of Figure 1). Lastly, the plain embellishment type was designed to involve no (or minimal) visual markup. The right-most column of Figure 1 shows examples of the plain variations.

The final independent variable was presence of color, which had two levels: color or black-and-white; the black-and-white variations were simply grayscale images of the original color versions of the charts. All charts were designed to be readable in both variations. For line and bar charts, this was not an issue, however, pie charts often use different colors to denote different data categories, so the most straightforward design for the plain variations of pie charts would have caused problems. Preliminary testing with grayscale revealed difficulty in distinguishing and matching different shades of gray; therefore, patterned regions were used to identify categories in the plain version of the chart type (see right-hand column of Figure 1). Generic patterns (e.g., striped, textured, dotted) were chosen in an effort to maintain the minimal aesthetic style required for the plain embellishment level. To test recall the experiment was designed as a two-part study, with one week in between study sessions. In the first session, participants viewed and completed simple data interpretation tasks. One week later, a the second session assessed participant's memory of those same charts.

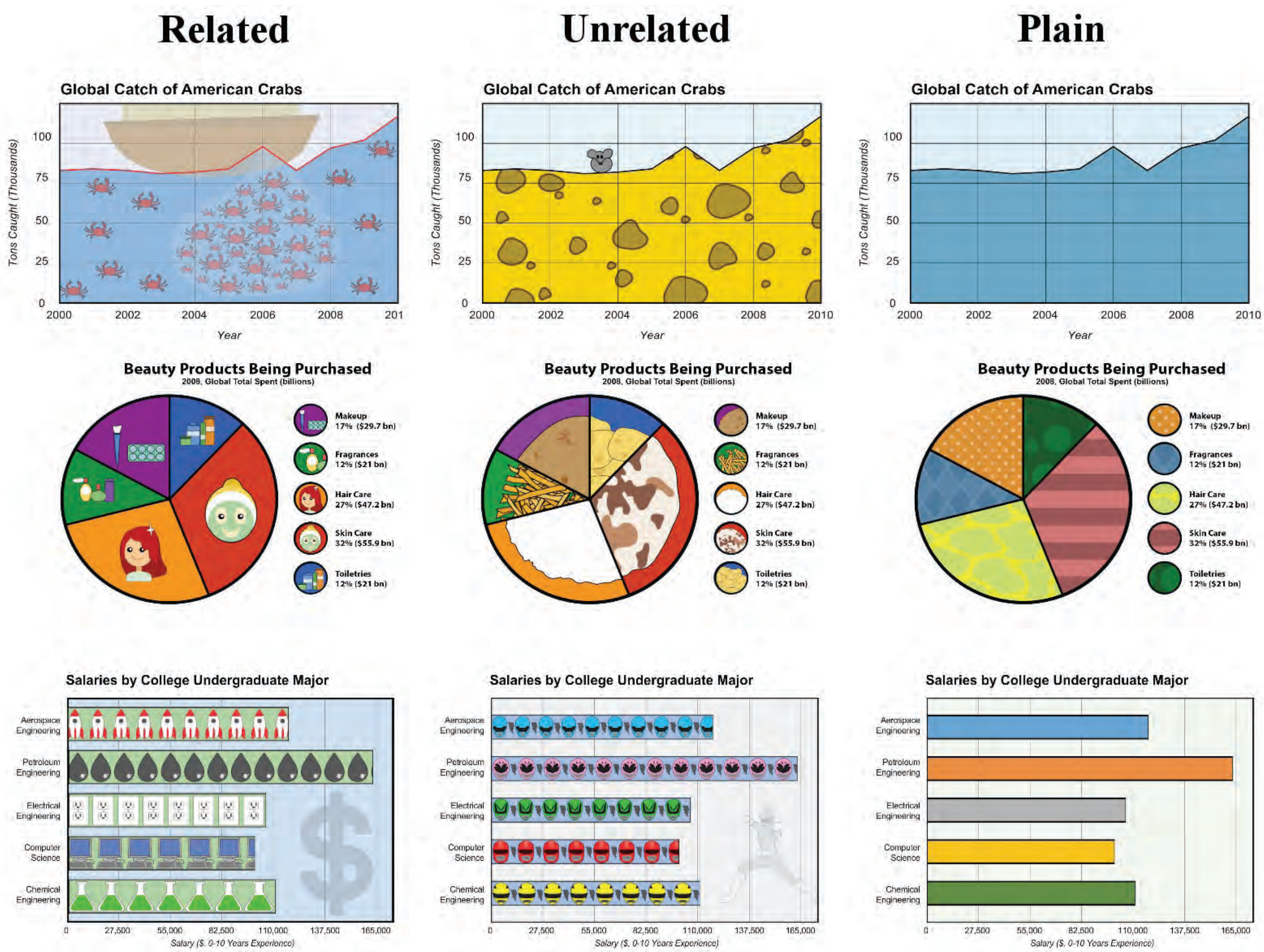

Figure 1 / Examples of charts used in the experiment. Each horizontal row consists of variations of the same chart modified according to the embellishment type factor. Each row also serves as an example of a different chart type for line, pie, and bar charts.

Dataset and Chart Development

In order to achieve experimental control for the different independent variables in the experiment we developed a dataset with a total of 180 charts. The charts were derived from 18 entries of base chart data divided evenly amongst the three chart types, each with five variations of embellishment types (two for unrelated, two for plain, and one for related). In total, we produced 90 charts (5 variants x 18 entries) to be used for each combination of the 3x3 within-subject factors. We then used the color charts to produce 90 more charts in greyscale, as color vs. black-and-white was controlled between subjects.

To support consistency, while also enabling creation of multiple unique combinations of charts, we developed design templates for the layout of bar, pie, and line chart types. Similarly, we developed template styles for the visual imagery of the embellishment types for each chart type. As seen in Figure 1, the template approach ensured that variations maintained an overall consistency in visual data encodings and chart design, and made it possible to change the embellishment imagery without modifying the underlying chart and data mappings. A single designer created all chart variations to maintain embellishment style across charts.

The data and themes for each chart were derived from the MASSVIS (Massachusetts (Massive) Visualization) dataset (Borkin et al., 2016) with numerous data visualizations scraped from online sources. Our criteria for choosing chart information was based on the complexity of the content, bias, understanding, and whether or not they could be represented visually. Occasionally, chart data was modified (e.g., title changes or minor value changes) to meet the purposes of the experiment. By creating our own visual imagery for the charts and a semi-automated system for developing the charts, we were able to maintain consistency in both style and layout.

The chart sets were generated to maintain similarity based on chart type. All pie charts had four or five categories, with all categories together totaling to 100%. Bar charts had five or six categories with varying ranges. Line charts had 9 – 24 data points based on how many were needed to show a trend, the ranges varied as well. When plotting values, we chose to make each chart with an axis range starting at zero in accordance with common visualization guidelines in order to reduce potential bias from deceptive visualizations (Pandey et al., 2015).

Tasks and Measures

The experiment consisted of two user-study sessions that included multiple question types. Participants were informed of both sessions ahead of time, but initial instructions were purposefully vague to avoid explicitly indicating the need to remember information across the sessions.

Session 1

In the first session, participants viewed 18 charts split by chart type and embellishment type. The initial ordering of the charts was randomized, as were whether the charts were in color or black and white. For each chart, two multiple choice questions were asked: identify and compare. The questions, refined in a pilot study, were designed to be readily understood, such that reviewing the chart would allow participants to determine a single correct answer from the multiple-choice options. The first question was always identification, which required participants to look up a certain value by interpreting the data graphic.

The following are examples of identification questions used in the study:

1. In 2008, which of the following beauty products was being purchased the most?
2. In 2011 what was the highest average number of hours spent in traffic in the U.S.?
3. Which location accounts for the least amount of robbery occurrences?

The second question was a comparison that required participants to interpret multiple data values and make a judgment about similarity or differences in magnitude. The following are examples of comparison questions from the study:

1. Which beauty product was purchased about as much as "Fragrances" in 2012?
2. In 2011, what month had a similar average

amount of hours spent in traffic to June?
3. Which location has the closest percentage of robbery occurrences to "Residences"?

When first viewing a complete chart with a title and label, the questions were not immediately visible. Participants were only shown the chart and asked to pay close attention to it. The questions were revealed either after five seconds or upon the participant clicking a button to signify they were ready. This brief delay was added to allow participants to familiarize themselves with the chart before they began reading questions and looking for solutions. The charts remained visible while they answered the multiple-choice questions, and a 'Confirm Answer' button was used to proceed to the next chart and set of questions. Exact time spent studying the visualizations varied by participant due to the flexibility of the online procedure, as the questions did not advance until all the correct answers were selected. This ensured that the users understood the chart contents. The system did not provide the correct answer when incorrect answers were given, and the rate of correct response was not included in analysis.

Session 2

The second session included four separate tasks: title recall, value recall, visual recall, and theme recall. As much as possible, these tasks and their order were designed so earlier tasks would not bias or assist recall for the following tasks. Table 1 provides an abbreviated summary and order of task types from session two. The title recall task was designed to assess how well participants could recall the titles of the charts viewed in the first session. It was administered as a recognition task, where no charts were visible, and participants were asked to select the titles of charts they recalled from the previous session. Titles were presented in a checkbox list of 30 randomly ordered titles distributed over three pages, 12 of which were fabricated. To promote the idea that some might be false, the participants were not required to select a title before proceeding to the next page.

For the value recall task, the goal was to test participants on their recall of the correct answers of the same identification and comparison multiple-choice questions that they answered in the first session. The ordering of the answer choices was the same as before, but charts were not available to assist in answering the questions.

The third task was visual recall, designed as an alternative measure of how well participants remembered the chart data values. Rather than requiring a numerical answer, the visual recall task was assessed based on the appearance, shape, or trend of the chart. Since session one included an identification question pertaining to a value on the chart, we tested visual memory of the charts themselves. For every chart presented in session one, participants were given an interactive version. To avoid interference with the following task, they did not include titles, labels, or axes. In addition, the data value from the identification question was modified from the original, in that the value was shown at half the chart's maximum value range. Participants were asked to adjust the value to match what was seen in session one (see Figure 2). The chart included an arrow that pointed to the adjustable data value—that is, a category for bar or pie charts, or a point for the line charts.

Participants used a pair of up-and-down arrow buttons to interactively increase or decrease the indicated value within a range; all values could be increased to the maximum of the value range and decreased to the minimum of the value range. The method for changing the data value and altering the visualization

Table 1 / A summary of evaluation tasks used to assess recall in Session 2 of the experiment.

Recall Task	Description
Title	From a list, select titles from session 1.
Value	Answer identify and compare questions from session 1 without chart.
Visual	Match the shape of provided chart to session 1 by interactively manipulating a value.
Theme	Given an unlabeled chart, select the theme best matching the charts from session 1.

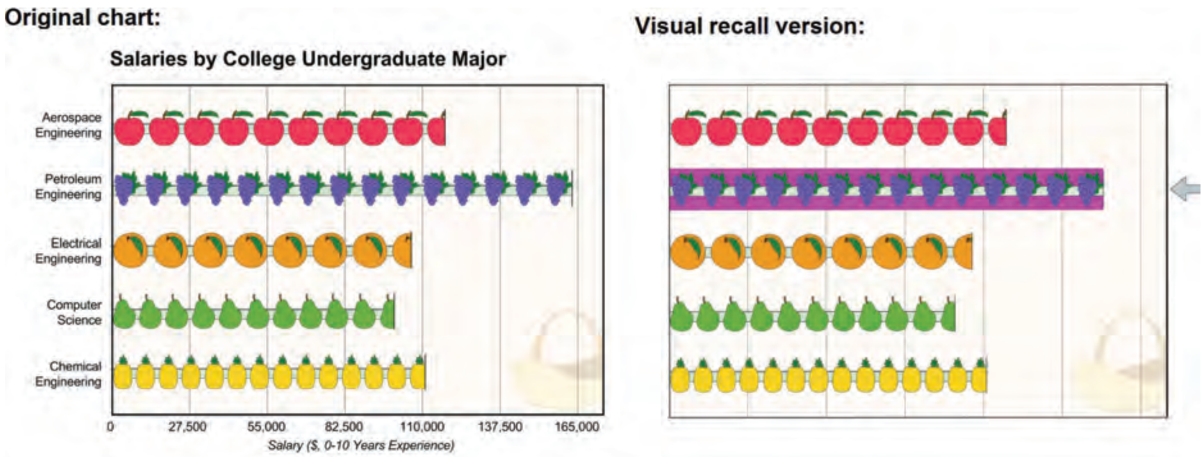

Figure 2 / In the visual recall task, participants interactively manipulated a value on the chart. The left image shows the original version of the chart seen in the first session, while the right shows the version used for the visual recall task of the second session. The arrow on the right side of the chart and the pink-highlighted background indicate the adjustable value. The interaction versions did not include textual labels.

varied based on appropriateness for the chart type. For instance, adjusting the value represented by a bar would increase/decrease the bar length without effecting other categories. However, adjusting any single value in pie or line charts affected the perception of neighboring values. To compensate, we also adjusted neighboring values to make any change compatible with the chart composition. When adjusting the value of a point in a line chart, the nearby points along the line were also adjusted using an exponential function to determine the weight of influence. Doing this allowed more realistic adjustments to the specified value without causing drastic spikes or dips in the chart and a similar approach was used for pie charts. To always keep the total of pie chart categories at 100%, changing one value required all other values to be inversely adjusted in an equal split.

Theme recall, the last task of session two, was used to evaluate how well participants could recall the theme of the chart if shown the chart without textual information. Participants were shown the unlabeled chart and asked to select the main theme from six choices. For example, if the chart was a pie chart of favorite Girl Scout cookies, the correct theme would be favorite cookies. The composition of incorrect options depended on the embellishment type of the given chart. If it was a related embellishment, the correct answer was in one of the three options related to the imagery, and the other three options concerned alternative topics. If embellished without context, three incorrect options were related to the imagery, two were fabricated as alternatives. For the plain conditions with no imagery, having multiple related answers could have been interpreted as hint about the nature of the correct answer; therefore, to avoid biasing participant answering, all five incorrect options were fabricated choices for the plain conditions.

Procedure

The study was conducted as an online study, with participants recruited through university email distribution and a ten-dollar (USD) participation incentive. Participants were not informed that they would be asked questions involving memory and recall; they only knew they would be answering questions about charts. Volunteers were sent a unique link to the web application. At the beginning of the first session, they were asked to digitally consent and fill out a background questionnaire. They were then shown instructions and an example of the types of multiple-choice questions they would be answering. The first session asked two questions for each of the 18 charts (a total of 36 questions) and had an average completion time of approximately 13 minutes.

Seven days later, participants were emailed with instructions for completing the second half of the study. To ensure similar amounts of time between sessions for participants, we required the second session to be completed within two days of the one-week notice, meaning that all successful participants completed the study within seven to nine days. Before each of the four tasks for this session (see Table 1), instructions with example images and questions were shown to clarify the task. The ordering of tasks was held consistent for all participants, though the ordering of charts was randomized for each participant. Completion of the second session took approximately 15 minutes.

Participants

The experiment had a total of 90 participants, but only 80 completed both sessions and did not identify as colorblind. Only participants who completed both sessions were included in the analysis. Of these, 32 identified as male, 47 identified as female, and 1 did not specify gender. Based on self-reports of education attainment, the majority of participants either completed high school (27) or held a Bachelor's (23) or Master's (23) degree. Several participants reported having PhD (4) or indicated "Other" (3). Ages ranged from 18 to 58, with a median age of 24.

RESULTS

We tested the effects of independent variables of embellishment type, chart type, and color on title, value, visual, and theme recall tasks from session two. Quantitative results from the tasks were shaped into a score by first assessing correctness per item for each specific task, then the scores were averaged for each task per participant. We analyzed the scores using 3x3x2 repeated-measure Analysis of Variance (ANOVA). If a significant main effect occurred, a Tukey test was performed for post-hoc analysis. We only discuss the significant effects in this report.

For title recall, where participants were asked to recall titles of charts shown in session one, the scores for the task were a count of the number of correctly selected titles. Figure 3 plots mean scores with standard error. There was a significant main effect in chart type, $F(2, 16) = 4.61$ and $p = 0.011$, and embellishment type, $F(2, 16) = 4.05$ and $p = 0.019$. Bar charts performed significantly better than line charts ($p < 0.01$). Related charts scored higher than unrelated charts ($p < 0.05$) and almost significantly better than plain charts ($p = 0.058$). Since this task was more concentrated on the message of the chart, it makes sense that related charts tended to score higher.

For value recall, where participants were asked to answer the same identify and compare questions from session one, the score was the average number of correct answers. There was a significant main effect in chart type with $F(2, 156) = 4.39$ and $p = 0.014$. Bar charts performed significantly better than line charts $p < 0.01$ (see Figure 4). It is worth noting that this same effect was significant for title recall. There were no effects of embellishments on recall.

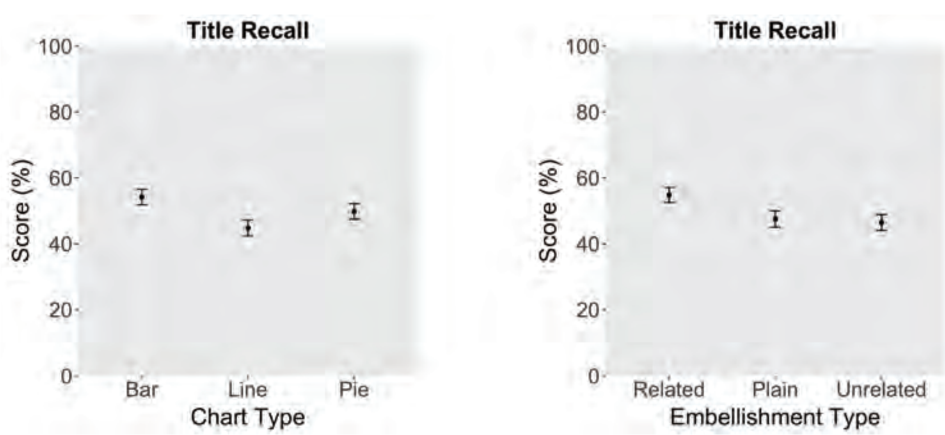

Figure 3 / Mean scores for title recall grouped by chart type (left) and embellishment type (right). Error bars show standard error. Bar charts scored significantly higher than line charts. The score of the related charts had significantly higher scores than unrelated charts and almost significantly higher scores than plain charts.

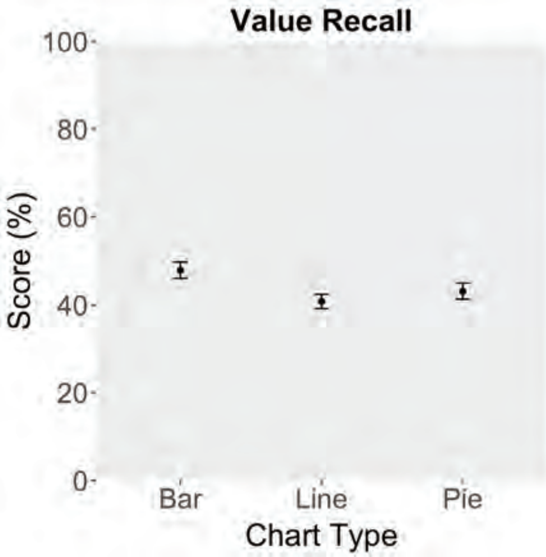

Figure 4 / Mean scores for value recall grouped by chart type. Error bars show standard error. Bar charts scored significantly higher than line charts.

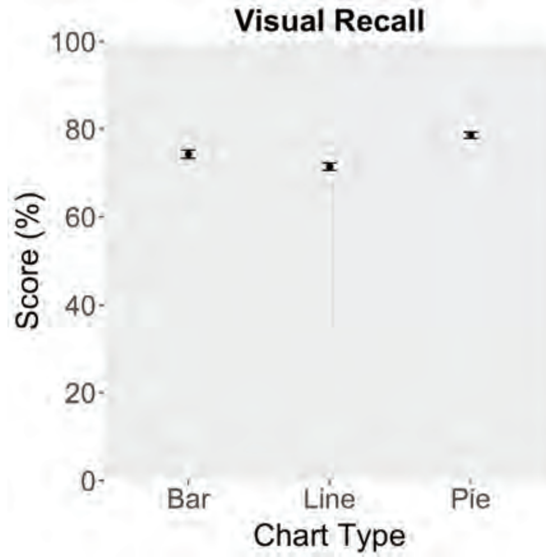

Figure 5 / Mean scores for visual recall grouped by chart type. Error bars show standard error. Pie charts scored significantly higher than bar and line charts. Bar charts scored significantly higher than line charts

For visual recall, participants were asked to manipulate a value of either a category or point on the chart to match their recall of the chart shape. We scored their final manipulation by taking the distance from the correct answer divided by the range of possible value to produce a value between zero and one. There was a significant main effect in chart type, with $F(2, 156) = 24.77$ and $p < 0.001$. Pie charts performed significantly better than both line charts and bar charts with $p < 0.001$ for both. Bar charts performed significantly better than line charts with $p < 0.05$ (see Figure 5). There were no effects of embellishments on recall.

For theme recall, participants were shown an unlabeled and untitled chart and asked to select the theme from a list of choices. The score for this task was a count of the number of correct selections. There was a significant effect in embellishment type with $F(2, 160) = 112.53$ and $p < 0.001$. The descending order ranking of the scores was embellished with context, plain, embellished without context (see Figure 6). Related charts scored more than plain and unrelated, both with $p < 0.001$, and plain had higher scores than unrelated with $p < 0.001$. This may be the result of participants recalling the imagery more than the actual subject.

There was a significant effect of chart type on the time it took to complete session one with $F(2, 156) = 3.31$ and $p = 0.009$. Overall, pie charts took more time to answer than bar charts ($p < 0.05$). The average times for title, value, visual, and theme recall tasks were 1.68, 5.94, 3.77, and 3.74 minutes, respectively. There was a

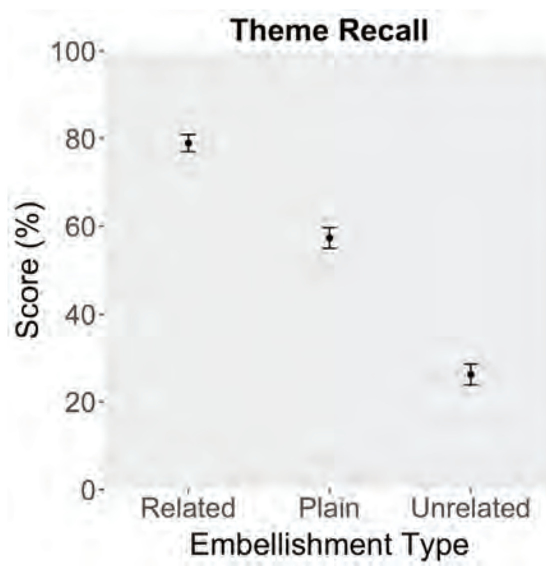

Figure 6 / Means scores for theme recall grouped by embellishment type. Error bars show standard error. Embellished charts scored significantly higher than plain and unrelated charts. Plain charts scored significantly higher than unrelated charts.

significant effect of chart type on the time it took to complete the theme recall task with $F(2, 156) = 3.71$ and $p = 0.027$, but no significant pairwise differences among chart types from post-hoc testing.

DISCUSSION

In this section, we discuss how the results relate with respect to the hypotheses, and we also discuss study limitations and implications for visualization creation tools.

Revisiting the Hypotheses

The results from the study can be used to evaluate our hypotheses described in the experiment section. We aimed to collect further empirical data about how different types of embellishments influence the memorability of different types of chart information. To help simplify the measures from the four tasks, we consider their organization into two main categories: chart topic (title recall, theme recall) and data details (value recall, visual recall), though these are not always clearly separate elements. We found that embellishment types had a significant effect on topics based on the results of title recall and theme recall scores. On the other hand, the study failed to detect differences for the data detail tasks. These results support hypothesis H1, as well as align with prior research on memorability (Bateman et al., 2010; Borkin et al., 2013; 2016). This finding also partially provides backing for the claim that embellishments can be effective for reinforcing a message about a data topic, though information about the message must be related to the data values, title, and theme. Further experimentation is needed to clearly evaluate memorability of the directionality of a message. It is not the case that the simple addition of any decorative elements will automatically increase memorability.

The opposite effect for relevant vs. unrelated embellishment is clear. Related embellishments significantly reinforced memory of chart topics, while unrelated embellishments negatively affected memorability. Thus, relevant embellishments can have a positive effect on memory, but only if the embellishments are chosen well. A probable explanation is that memorability of the embellishments overrides the memorability of the actual informational content. This indicates that a poor choice for supplemental embellishment ("chart junk") can be distracting or interfere with memory if the viewer does not perceive a relationship between embellishments and topic. Furthermore, this finding establishes a baseline for future studies, which may systematically vary facets of relevance with respect to design elements in a given visualization.

Regarding hypothesis H3, the differences between color and black-and-white images was clearly minor in our chart set and evaluation. There was no significant effect of color type on charts; thus, the results do not contribute evidence in support of H3. Our study did not account for the use of appropriate or unusual coloration in relation to the embellishment design, which would also be interesting to consider along with the study of relevance.

We also found that chart type played a role in memorability. For all tasks except theme recall, chart type had significant effects on recall. These results provide evidence in support of hypothesis H2 by suggesting that chart type is influential for memory of chart information. However, there was little consistency with which chart type (line, pie, and bar) were more or less memorable for the different recall measures. For example, bar charts were significantly more memorable than line charts for title recall, value recall, and visual recall. Pie charts were more memorable than both line charts and bar charts in terms of visual recall. The high memorability of pie and bar charts for visual recall might suggest that viewers had an easier time remembering visual ratios, as that is largely what pie and bar charts represent, rather than line shape or trends. The lack of effect of chart type on theme recall may be similar to prior findings that common chart types may be less memorable than unique visualizations (Borkin et al., 2013). In other words, because the experiment only included three basic chart types, their designs were not interesting enough to impact thematic memory.

Towards a Greater Understanding of Memorability and the Designers' Role

A number of studies have addressed the topic of memorability in infographics and embellished charts.

While some of these studies (Bateman et al., 2010; Borkin et al., 2013; Li and Moacdieh, 2014) look at illustrative graphics and elaborate infographic embellishments more similar to Holmes (1984), others have investigated more minor manipulations and stylistic enhancements (Borgo et al., 2012; Haroz et al., 2015). Our research tends toward the side of more limited embellishments in an attempt to preserve the underlying visual encodings of data, though our custom created embellishments may provide a somewhat greater degree of enhancement while maintaining high experimental control for the presented study. Overall, the complexity of the design space and range of variations in study designs can make it difficult to directly compare the results of different investigations on the topic of memorability.

One of the primary contributions of our research is the controlled comparison of fundamental chart types. While other studies have considered a variety of chart types from existing available infographics, our method and hand-crafted data set allowed a more controlled approach. The results demonstrate greater visual memorability of pie charts compared to bar and line charts, whereas bar charts facilitated better memory of information values. This is an important finding that might suggest people fundamentally allocate attention differently when interpreting different chart designs. Perhaps because bar charts support more accurate perceptual judgments due to the use of length encodings (Cleveland and McGill, 1984; Heer and Bostock, 2010; Saket et al., 2018), participants were more easily able to extract the value details for retention. Alternatively, perhaps pie charts are viewed more informally, or they may encourage more attention to the visuals holistically rather than to precise judgments, which may explain why pie charts are more visually memorable. As another explanation, it may be possible that pie charts or bar charts were easier to visually interpret during the visual recall measure rather than the representation itself necessarily being more memorable; further research is needed for a deeper understanding of possible factors.

Our experiment assesses recall after a significantly longer period of time than some memorability studies such as Borgo et al. (2012) and Haroz et al. (2015). Our study, with one week between sessions, is more similar to Bateman et al.'s (2010) approach of 2 – 3 weeks between sessions. Similar to our findings, Bateman et al. investigated a variety of aspects of chart understanding, but no evidence that embellishment were better for memory of information beyond the designer's primary intended message was found. This result is most similar to our findings for theme recall, though we also contribute further information about relevance about embellishments, and we also found effects for title recall.

The study by Haroz et al. (2015) resulted in partially similar findings regarding relevance of embellishments to the data topic, but their study's emphasis on "close-but-not-quite-matching" embellishments offers the potential for a different type of distraction or confusion if viewers mix up specific data items with the embellishment. Our study's use of related and unrelated embellishment takes a more extreme approach to verify that the unrelated versions were clearly not related to the topic. Importantly, the negative results of unrelated embellishments hold true across studies, suggesting that the effects observed by Haroz et al. were likely not due to interference from "class similarity." Together, our results suggest the reduced memorability of chart information may be due to overall distraction and difficulty in separating the visuals from the data, and our study demonstrates that these the negative effect of unrelated visuals persisted over a one week period.

Following the existing body of research, our study found that chart design can affect memorability, but different studies evaluate different aspects of memorability and chart information. Our separation of value recall, title recall, visual recall, and theme recall tasks help separate different types of information. We can examine consistencies by cross-referencing studies. Borkin et al. (2016) found that presence of a chart title contributes to greater memorability in general, and our results demonstrate related embellishments did improve title retention. The assessment of "concept grasping" by Borgo et al. (2012) was similar to our assessment of theme recall, and the findings align as previously discussed. Recall of chart value details is difficult, and our study did not find differences due to the embellishment conditions. In contrast, Haroz et al. (2015) did observe significant differences in value recall from embellishments, but their tested visual designs involved limited numerical ranges and discrete values shown by a corresponding number of representative icons. It may be the case that such representations are more visually memorable, as perhaps participants were recalling the chart appearance rather than completely encoding the data values. Given the shorter period between exposure and recall, we expect a visual recall strategy may be feasible. In our own study, participants demonstrated fairly high performance on the visual recall task even one week later, but it was not sufficient for supporting recall of details. Future studies may examine the possibility that participants are encoding shape or relational characteristics of the visualization, and how design choices may impact such encodings.

Implications for Graphic Creation Tools

The results of this paper suggest that design enhancements alone are not necessarily sufficient to positively influence human-centered measures such as memorability. These results may therefore hold implications for the feature sets of graphic authoring tools. Specifically, in the information visualization community, multiple recent initiatives have led to significant advances in visualization creation and authoring tools (Kim et al., 2018; Liu et al., 2018; Ren et al., 2018; Xia et al., 2018). For example, Liu et al. (2018) developed Data Illustrator, a software that combines vector editing tools like Adobe Illustrator with data-bindings to enable the creation of complex, custom visuals. The results of this research may imply that such functionality could be extended to ensure that users have the ability to craft embellishments that align with their particular data domain. More broadly, the recent increase in visualization authoring advancements may point to a need for more robust experimental methodologies. As more creative representations of data become possible, quantitative evidence is needed to determine whether particular compositions of visualization styles and embellishments lead to positive outcomes for the audience.

LIMITATIONS

In this study, we used base data from the MASSVIS dataset and modified the visual representations to meet the needs of the experiment. Deviating from the original design of the infographics and visualizations altered the style, tone, and message of the source. As such, our experiment does not claim to use real infographics or informational signage; this was intentional and necessary to maintain experimental control of our independent variables. Real infographics are often more complex and involve additional types of embellishment beyond the templated decoration formats included in our study. Embellishments in our experiment were limited to additions of imagery within and around the visual data encodings of each chart. Moreover, since we purposefully maintained the mapping between visual encodings and data values across conditions, our experiment did not consider more advanced types of stylistic manipulations that would have altered the data presentation or "lie factor" (Tufte, 1983). Such manipulation or deception was beyond the scope of this research, though the topic has been investigated by others (see Correll and Heer, 2017a; Pandey et al., 2015). Further research is needed to gain a broader understanding of memorability of visuals and information across a variety of media formats, such as web, signage and print, in the wild.

CONCLUSION

Our study tests the effects of graphical enhancements, color, and chart type on memory of chart information over the period of a week. For this particular study, the enhancements were supplemental embellishments that did not manipulate the fundamental chart design or data encodings. Embellishments relevant to the data significantly helped in the recall of title and thematic elements, and unrelated embellishments were significantly detrimental to memorability. Having this knowledge, we can exploit more effective ways in conveying both the message and visual displays. Future work might include understanding the effects of interactions and embellishments that alter data values to strengthen the communication of an intended message.

REFERENCES

Atkinson, R.C. and Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. *Psychology of Learning and Motivation 2*, 89–195.

Bateman, S., Mandryk, R.L., Gutwin, C., Genest, A., McDine, D. and Brooks, C. (2010). Useful junk?: The effects of visual embellishment on comprehension and memorability of charts. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2573–2582.

Beecham, R., Dykes, J., Meulemans, W., Slingsby, A., Turkay, C. and Wood, J. (2017). Map LineUps: Effects of spatial structure on graphical inference. *IEEE Transactions on Visualization and Computer Graphics* 23(1), 391–400.

Borgo, R., Abdul-Rahman, A., Mohamed, F., Grant, P.W., Reppa, I., Floridi, L. and Chen, M. (2012). An empirical study on using visual embellishments in visualization. *IEEE Transactions on Visualization & Computer Graphics* 18, 2759–2768.

Borkin, M.A., Bylinskii, Z., Kim, N.W., Bainbridge, C.M., Yeh, C.S., Borkin, D., Pfister, H. and Oliva, A. (2016). Beyond memorability: Visualization recognition and recall. *IEEE Transactions on Visualization and Computer Graphics* 22(1), 519–528.

Borkin, M.A., Vo, A.A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A. and Pfister, H. (2013). What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics* 19(12), 2306–2315.

Bylinskii, Z., Kim, N.W., O'Donovan, P., Alsheikh, S., Madan, S., Pfister, H., Durand, F., Russell, B. and Hertzmann, A. (2017). Learning visual importance for graphic designs and data visualizations. *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, 57–69.

Chase, W.G. and Simon, H.A. (1973). Perception in chess. *Cognitive Psychology* 4(1), 55–81.

Claes, S. and Vande Moere, A. (2013). Street infographics: Raising awareness of local issues through a situated urban visualization. *Proceedings of the 2nd ACM International Symposium on Pervasive Displays*, 133–138.

Cleveland, W.S. and McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association* 79(387), 531–554.

Correll, M. and Heer, J. (2017a). "Black hat visualization." Paper presented at DECISIVe Workshop, IEEE VIS 2017, Phoenix, Oct 2.

Correll, M. and Heer, J. (2017b). Regression by Eye: Estimating Trends in Bivariate Visualizations. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1387–1396.

Craik, F. I. and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General* 104(3), 268.

Few, S. (2011). "The chartjunk debate: A close examination of recent findings." *Perceptual Edge,* April/May/June 2011. http://www.perceptualedge.com/articles/visual_business_ intelligence/the_chartjunk_debate.pdf

Haroz, S., Kosara, R. and Franconeri, S.L. (2015). Isotype visualization: Working memory, performance, and engagement with pictographs. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 1191–1200.

Harrison, L., Yang, F., Franconeri, S. and Chang, R. (2014). Ranking Visualizations of Correlation Using Weber's Law. *IEEE Transactions on Visualization and Computer Graphics* 20(12), 1943–1952.

Heer, J. and Bostock, M. (2010). Crowdsourcing graphical perception: Using mechanical turk to assess visualization design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 203–212.*

Holmes, N. (1984). *Designer's guide to creating charts & diagrams.* New York: Watson-Guptill.

Jardine, N., Ondov, B.D., Elmqvist, N. and Franconeri, S. (2019). The Perceptual Proxies of Visual Comparison. *IEEE Transactions on Visualization and Computer Graphics* 26(1), 1012–1021.

Kellaris, J.J. and Machleit, K.A. (2016). *Signage as Marketing Communication: A Conceptual Model and Research Propositions. Interdisciplinary Journal of Signage and Wayfinding* 1(1).

Kim, N., Schweickart, E., Liu, Z., Dontcheva, M., Li, W., Popovic, J. and Pfister, H. (2018). Data-driven guides: Supporting expressive design for information graphics. *IEEE Transactions on Visualization & Computer Graphics* 23(1), 491-500.

Kosara, R. (2007). Visualization criticism: The missing link between information visualization and art. *11th International Conference Information Visualization*, 631–636.

Lahrache, S., El Ouazzani, R. and El Qadi, A. (2018). Visualizations memorability through visual attention and image features. *Procedia Computer Science* 127, 328–335.

Li, H. and Moacdieh, N. (2014). Is "chart junk" useful? An extended examination of visual embellishment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 58(1), 1516–1520.

Lin, S., Fortuna, J., Kulkarni, C., Stone, M. and Heer, J. (2013). Selecting semantically-resonant colors for data visualization. *Computer Graphics Forum* 32, 401–410.

Liu, Z., Thompson, J., Wilson, A., Dontcheva, M., Delorey, J., Grigg, S., Kerr, B. and Stasko, J. (2018). Data Illustrator: Augmenting Vector Design Tools with Lazy Data Binding for Expressive Visualization Authoring. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.*

Nuthall, G. (2000). The role of memory in the acquisition and retention of knowledge in science and social studies units. *Cognition and Instruction* 18(1), 83–139.

Pandey, A.V., Rall, K., Satterthwaite, M.L., Nov, O. and Bertini, E. (2015). How deceptive are deceptive visualizations?: An empirical analysis of common distortion techniques. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 1469–1478.

Quispel, A., Maes, A. and Schilperoord, J. (2016). Graph and chart aesthetics for experts and laymen in design: The role of familiarity and perceived ease of use. *Information Visualization* 15(3), 238–252.

Ren, D., Lee, B. and Brehmer, M. (2018). Charticulator: Interactive Construction of Bespoke Chart Layouts. *IEEE Transactions on Visualization and Computer Graphics* 25(1), 789-799.

Rensink, R.A. and Baldridge, G. (2010). The perception of correlation in scatterplots. *Computer Graphics Forum* 29, 1203–1210.

Ritchie, D, and Karge, B.D. (1996). Making information memorable: Enhanced knowledge retention and recall through the elaboration process. Preventing School Failure: *Alternative Education for Children and Youth* 41(1), 28–33.

Ruchikachorn, P. and Mueller, K. (2015). Learning visualizations by analogy: Promoting visual literacy through visualization morphing. *IEEE Transactions on Visualization and Computer Graphics* 21(9), 1028–1044.

Saket, B., Srinivasan, A., Ragan, E.D. and Endert, A. (2018). Evaluating interactive graphical encodings for data visualization. *IEEE Transactions on Visualization and Computer Graphics* 24(3), 1316–1330.

Skog, T., Ljungblad, S. and Holmquist, L. E. (2003). Between aesthetics and utility: Designing ambient information visualizations. *IEEE Symposium on Information Visualization*, 233–240.

Smith, D.A. and Graesser, A.C. (1981). Memory for actions in scripted activities as a function of typicality, retention interval, and retrieval task. *Memory & Cognition* 9(6), 550–559.

Szafir, D.A. (2017). Modeling Color Difference for Visualization Design. *IEEE Transactions on Visualization and Computer Graphics 24(1), 392-401.*

Tufte, E. (1983). *The visual display of quantitative information*. Cheshire: Graphics Press.

Wickham, H., Cook, D., Hofmann, H. and Buja, A. (2010). Graphical inference for infovis. *IEEE Transactions on Visualization and Computer Graphics* 16(6), 973–979.

Xia, H., Henry Riche, N., Chevalier, F., De Araujo, B. and Wigdor, D. (2018). DataInk: Direct and Creative Data-Oriented Drawing. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.*

Zacks, J., Levy, E., Tversky, B. and Schiano, D.J. (1998). Reading bar graphs: Effects of extraneous depth cues and graphical context. *Journal of Experimental Psychology: Applied* 4(2), 119.