Psychology of Aesthetics, Creativity, and the Arts

Impressionism-Inspired Data Visualizations Are Both Functional and Liked

Pavel Kozik, Laura G. Tateosian, Christopher G. Healey, and James T. Enns

Online First Publication, May 14, 2018. http://dx.doi.org/10.1037/aca0000175

American Psychological Association
Psychology of Aesthetics, Creativity, and the Arts

1931-3896/18/$12.00
http://dx.doi.org/10.1037/aca0000175

Impressionism-Inspired Data Visualizations Are Both Functional and Liked

Pavel Kozik
University of British Columbia

Laura G. Tateosian and Christopher G. Healey
North Carolina State University

James T. Enns
University of British Columbia

Creating data visualizations that are functional and aesthetically pleasing is important yet difficult. Here we ask whether creating visualizations using the painterly techniques of impressionist-era artists may help. In two experiments we rendered weather data from the Intergovernmental Panel on Climate Change into a common visualization style, *glyph*, and impressionism-inspired painting styles, *sculptural*, *containment*, and *impasto*. Experiment 1 tested participants' recognition memory for these visualizations and found that impasto, a style resembling paintings like *Starry Night* (1889) by Vincent van Gogh, was comparable with glyphs and superior to the other impressionist styles. Experiment 2 tested participants' ability to report the prevalence of the color blue (representative of a single weather condition) within each visualization, and here impasto was superior to glyphs and the other impressionist styles. Questionnaires administered at study completion revealed that styles participants liked had higher task performance relative to less liked styles. Incidental eye tracking in both studies also found impressionist visualizations elicited greater visual exploration than glyphs. These results offer a proof-of-concept that the painterly techniques of impressionism, and particularly those of the impasto style, can create visualizations that are functional, liked, and encourage visual exploration.

Keywords: aesthetics, impressionism, information visualization

Supplemental materials: http://dx.doi.org/10.1037/aca0000175.supp

Communicating data effectively is one of the most important tasks of science. Although recent technological innovation has made large complex data sets more numerous and widely accessible than ever before, access in itself does not lead to better understanding. Information visualization, a branch of computer science, tackles the problems inherent in understanding such data sets (Card, Mackinlay, & Shneiderman, 1999; Healey, Tateosian, Enns, & Remple, 2004). Here large-scale spatial data sets, as found in weather, biomedical, and sport science research, are rendered into two-dimensional images where visual properties like color, size and orientation represent data. For example, in weather data, these visual properties might be used to portray temperature, wind speed, and wind direction (Joshi, Caban, Rheingans, & Sparling, 2009; Müller, Reihs, Zatloukal, & Holzinger, 2014; Pileggi, Stolper, Boyle, & Stasko, 2012; Ware & Plumlee, 2013). By doing so it is hoped data will be well represented, summarized, and that the viewer will detect trends and patterns among the variables that may otherwise have gone unnoticed.

Glyph-based representations are among the most commonly used data visualizations (Borgo et al., 2012, 2013). These visualizations consist of numerous visual elements, called *glyphs*, whose visual properties are determined by the underlying data they represent at a given spatial location. For example, by examining how glyph properties (e.g., relative shape, size, color, orientation, density, saturation, fuzziness and transparency) change across spatial locations users can understand how geographic areas or anatomical locations of the human body differ from one another (Ropinski, Oeltze, & Preim, 2011). Yet, despite their strong potential, researchers have begun to question their functionality (i.e., their memorability and ability to accurately convey data trends) and their aesthetic appeal (Borkin et al., 2013; Chen, 2005; Lee, Butavicius, & Reilly, 2003; Ward, 2008).

An alternative approach that combines these two concerns begins with the premise that beauty and functionality are intrinsically intertwined, such that what is beautiful is often useful and what is useful is often beautiful (Borkin et al., 2013; Chen, 2005; Healey & Enns, 2012; Lau & Moere, 2007; Norman, 2005; Tateosian, Healey, & Enns, 2007). Support for this perspective can be found in advertising and product design research showing that high

Pavel Kozik, Department of Psychology, University of British Columbia; Laura G. Tateosian and Christopher G. Healey, Department of Computer Science, North Carolina State University; James T. Enns, Department of Psychology, University of British Columbia.

Correspondence concerning this article should be addressed to Pavel Kozik, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, British Columbia, Canada, V6T 1Z4. E-mail: pavelkozik@psych.ubc.ca

ratings of aesthetic quality associate with greater visual exploration and perceived usability (Maughan, Gutnikov, & Stevens, 2007; Tractinsky, Katz, & Ikar, 2000). By adopting this approach, the present study seeks to systematically examine the relationship between visualization functionality and aesthetic appeal. We do so by creating data visualizations through the painterly techniques of prominent impressionist era artists (Brown, 1978; Cutting, 2006; Schapiro, 1997). To further ground this work in current visualization efforts, we compared three impressionism-inspired visualization styles with that of glyphs in two experiments.

In Experiment 1 the four visualization styles were tested for their short-term memorability (using a new-old recognition task), and in Experiment 2 the same styles were compared in their ability to accurately convey data frequency (using a numerosity estimation task). In both experiments, the aesthetic quality and perceived functionality of the visualizations were measured via questionnaires, and their influence on visual exploration assessed by incidental eye-tracking. Before describing the experiments in detail however, we first describe the impressionist techniques used to generate these visualizations.

Impressionist Painting Techniques

Talented artists rely on a variety of painterly techniques to engage and orient viewer attention, ranging from those that operate at the global level of the image to those that operate at the local level of individual brushstrokes (Hogarth, 1753; Kirby et al., 2003; Koenderink, van Doorn, & Wagemans, 2012; van Gogh, 1937). For instance, artists paint certain portions of an image to have greater detail and others less detail, thereby implicitly guiding the viewer (DiPaola, Riebe, & Enns, 2010, 2013). Artists may also distort visual properties like shape and color to further selectively bias the viewer toward some objects and features over others (Cavanagh, 2005; Ramachandran & Hirstein, 1999). Even low-level features, such as the artist's choice of dominant edge orientation, may be biased by a human viewer tendency to favor horizontal and vertical edges over oblique edges (Latto, Brain, & Kelly, 2000; Latto & Russell-Duff, 2002).

Within this literature, Tateosian and colleagues (2007) examined various well-known impressionist-era painters such as Claude Monet, George Seurat, and Vincent van Gogh. Analyses of their artworks revealed numerous but consistent differences in brushwork technique, including variation in paint thickness, stroke curvature, and color variegation. Parallels were also reported between the artistic techniques used by these artists and visual features known to be rapidly perceived by the human visual system. This prompted the authors to identify three distinct styles titled *interpretational complexity*, *indication and detail*, and *visual complexity*. They went on to then create algorithms that mimicked the main features of these styles. The present study uses these three styles but here we replace the original labels with *sculptural*, *containment*, and *impasto*, respectively, to refer to a core feature of each style. An illustration of these styles and their defining techniques is shown in Figure 1, where identical weather data have been rendered as a glyph visualization (Figure 1A), and then in each of the three impressionism-inspired styles (Figure 1B, 1C, 1D).

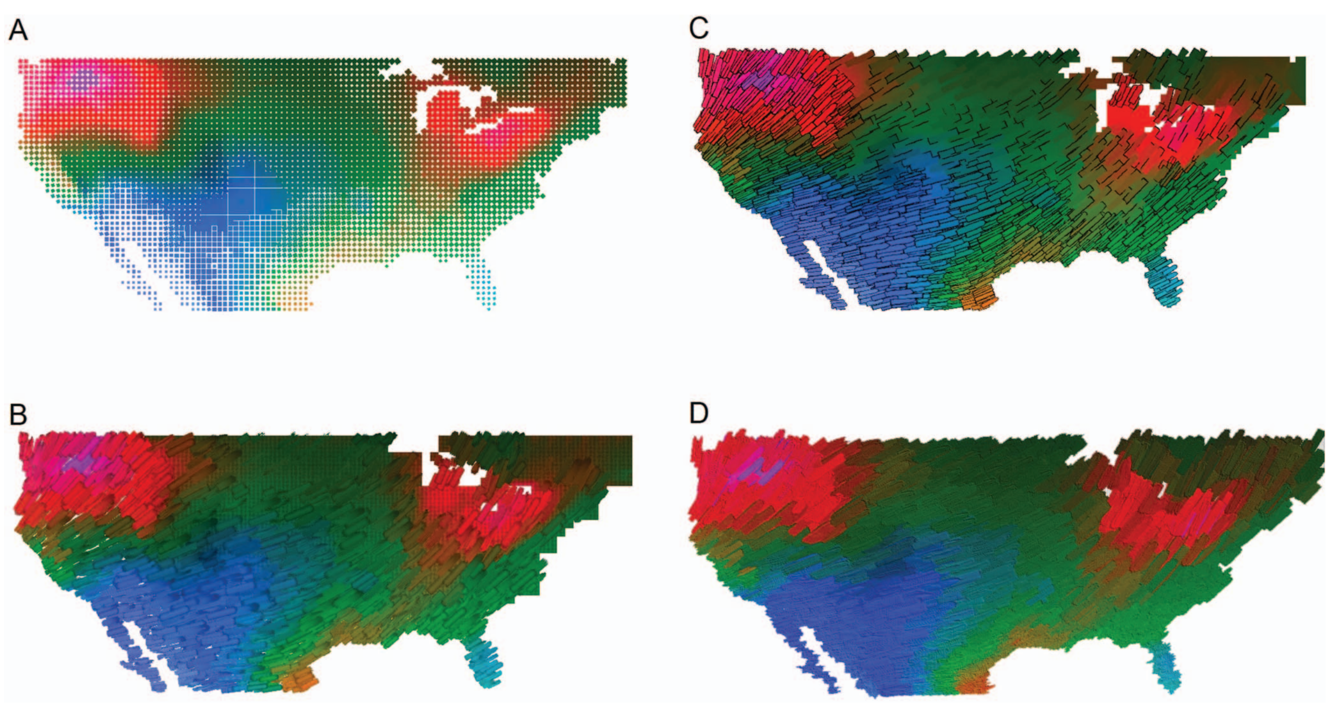

Figure 1. An example of weather data rendered into the four visualization styles: (A) glyphs, (B) sculptural, (C) containment, (D) impasto. In this example the weather properties of cloud cover, mean temperature, frost frequency, and wet day frequency are, respectively, represented by color, greyscale, size, and orientation. See the online article for the color version of this figure.

Several differences in visualization style are readily visible when viewing Figure 1. The distinct glyph visualization is composed of several nonoverlapping rectilinear shapes whose visual properties of color, greyscale, size and orientation are determined by dataset values like cloud cover, mean temperature, frost frequency and wet day frequency. The three impressionist styles, by way of comparison, are instead composed of overlapping brushstrokes. Considering each impressionist style in turn, sculptural (Figure 1B) mimics an oil painting approach. The data are first filtered at a global level to detect regions of rapid change and so the first layer becomes an undercoat that broadly defines the global shape and outline of the visualization. Finer details are then rendered in a subsequent layer through individual brushstrokes that provide detail only in regions of rapid change. Where data are highly homogenous results in an undercoat canvas that is plain in appearance. As such this style seeks to benefit from relative simplicity (Pezdek & Chen, 1982; Pezdek et al., 1988) as the viewer's attention is guided to the more detailed data regions of the image in a subtle and nonintrusive manner (Ramachandran & Hirstein, 1999). Visualizations in this style borrow heavily from paintings like *Water Lilies, Evening Effect* (1897–1899) by Claude Monet.

In the style called containment (Figure 1C) areas of high data variation are emphasized more deliberately to the viewer. Here brushstrokes defined by high-contrast outlines are applied to areas of high data variation whereas lower-contrast brushstrokes are applied in areas of low data variation. Partially outlined brushstrokes transition between these two zones under the assumption that the viewer's attention will be signaled toward the regions of greatest detail (DiPaola et al., 2010; Ramachandran & Hirstein, 1999). This style resembles various currently used pen-and-paper visualizations techniques (Lu et al., 2002), such as the "Sketchy" style in Wood et al. (2012) that participants found engaging and encouraging of free-hand note taking. Visualizations in this style reference Art Nouveau designs in which artists, such as Alphonse Mucha, were strongly influenced by the Impressionist movement.

The final impressionist style called *impasto* (Figure 1D) goes one step further in trying to capture data properties by using within-brush stroke variation as sometimes used by masterful artists. Here a single brushstroke, from start to finish, may change in color, paint thickness, contour, curvature, length, and orientation, based on the underlying dataset values. This style borrows heavily from Vincent van Gogh (1937) who wrote "I should like to paint in such a way that everybody, at least everybody with eyes, would see it. . . . I am endeavoring to find a brush work that is nothing but the varied stroke." As implied in the label, this style relies heavily on the technique of *impasto,* in which a thick overlapping of paint implies image texture. By combining these different techniques portions of the visualization appear almost decorative and the viewers' attention is attracted by both global and local image features. Support for the effectiveness and liking of this style comes from studies showing that data depictions with a decorative appearance are accurately perceived, aid memory performance, and are rated as aesthetically pleasing (Bateman et al., 2010; Borgo et al., 2012).

Considering these three impressionist styles in light of Vincent van Gogh's work allows for further contextualization and comparison. The techniques of sculptural borrow from artworks like *Starry Night Over the Rhone* (1888), whereas the techniques used in *Irises* (1889) resemble containment. The final impressionist style, impasto, is inspired from van Gogh's most famous work, *Starry Night* (1889). If van Gogh's timeline of experimentation, and the public's subsequent appreciation of it are to be considered, then one provocative (albeit speculative) hypothesis is that the style van Gogh is best known for today is also the visualization technique that will result in the most memorable and expressive impressionist style for modern users of data.

Overview of the Study

Source data sets for the present study were obtained from the Intergovernmental Panel on Climate Change (2014) and were rendered in the four different visualization styles (described above and in further detail in Tateosian et al., 2007). Four distinct geographic areas were chosen to diversify weather trends and the overall shape of visualizations, so that our interpretation of the results would not be limited to a specific region and monthly pattern. We had no a priori hypothesis concerning geographic area and so any differences reflecting them were treated as exploratory. Weather data in particular are ideal for experimental purposes as they are multivariate, have real-world immediate implications, are likely to already be somewhat familiar to participants, and frequently belong to a category of data sets, government, and news media, for which visualization memorability often suffers and research hence applicable (Borkin et al., 2013; Fabrikant, Hespanha, & Hegarty, 2010).

In Experiment 1 participants completed a series of trial blocks, where visualizations were rated on their complexity and arousal before a memory test was completed. In Experiment 2 participants viewed visualizations and were then asked "What percentage of the previous map was blue?", where blue represented a single weather condition. Color was selected as the visual property to estimate because color is critical to forming visualization first impressions (Harrison, Reinecke, & Chang, 2015) and is an effective means to guide viewer attention (Borgo et al., 2013; Fabrikant, Christophe, Papastefanou, & Maggi, 2012). At the end of both experiments participants completed a questionnaire in which they were asked which style they most liked and thought was most memorable (Experiment 1) or best for identifying the color blue (Experiment 2). Visual exploration in both experiments was measured via a noninvasive infrared eye tracker.

The primary study question was if impressionist styles would have greater functionality than glyphs. Functionality was defined in Experiment 1 as higher recognition accuracy and in Experiment 2 as more accurate numerosity estimation. The hypothesis we favored was that impasto would be particularly effective in comparison to glyphs and the other impressionist styles because it combined characteristics to attract attention based on both global factors (large scale variations in relative brushwork detail and the use of impasto) and local factors (greater heterogeneity of features within-brushstroke) and borrows from the painterly techniques associated with van Gogh's most famous works.

A secondary hypothesis was to examine the link between participants' subjective experience and their objective performance (Norman, 2005). In both experiments, subjective liking was measured by asking participants which style they most liked. In Experiment 1 subjective impressions of functionality were assessed by asking participants which visualization style they thought was the most memorable and in Experiment 2 which style they thought was most effective for color estimation. We hypothesized that impressionist visualizations would be more liked and would elicit greater visual exploration (Maughan et al., 2007) as they were created by harnessing the expertise of masterful artists.

Experiment 1: Memorability

An important attribute of an effective visualization is memorability. Experiment 1 compared glyph and impressionist-inspired visualizations on a new-old recognition task. The top panel of Figure 2 provides a visual overview of the experimental design. Experiments were approved by the University of British Columbia's Behavioral Research Ethics Board (H14-00037).

Method

Participants. Thirty undergraduates (15 males and 15 females, *mean age* = 20.20 *SD* = 2.01) were recruited through a voluntary university human subject pool and participated in exchange for course credit. All participants provided written informed consent, had normal or corrected-to-normal vision, were naive to the experimental hypothesis, and were debriefed upon study completion. All participants finished the study within one hour.

Stimuli and apparatus. Visualizations were generated based on data provided by the Intergovernmental Panel on Climate Change (Intergovernmental Panel on Climate Change, 2014). The visualizations consisted of 32 weather data sets from the years 1961 to 1990. Each visualization depicted a wide variety of weather conditions across the geographic areas of Africa, Asia, South America, and the United States, including monthly mean temperature, precipitation, wind speed, cloud cover, frost frequency, vapor pressure, and solar radiation. Each of the 32 weather data sets were rendered with all four visualization styles: glyphs, sculptural, containment, and impasto (Tateosian, 2006; Tateosian et al., 2007). This produced a total set of 128 visualizations. Visualizations were displayed on a 20.2 × 17.3 in. LCD monitor, set to a refresh rate of 60-Hz with a screen resolution of 1600 × 1200 pixels. Each visualization was screen centered on top an intermediate gray background at a resolution of 750 × 750 pixels. Because visualizations are frequently encountered on a similarly sized computer screen these resolution settings were comparable to what participants might encounter outside of the laboratory. A desktop computer controlled the presentation of instructions, practice, rating and memory trials as well as recorded participants' responses. Participants' eye-fixations were recorded throughout each testing session by an SR Eyelink 1000 set to a sampling rate of 1500Hz.

Procedure. Participants were welcomed to the laboratory, and after providing written study consent, were seated in a chair and asked to rest their heads on a chinrest 70 cm from the monitor display. The experimenter was present in the room for the entire

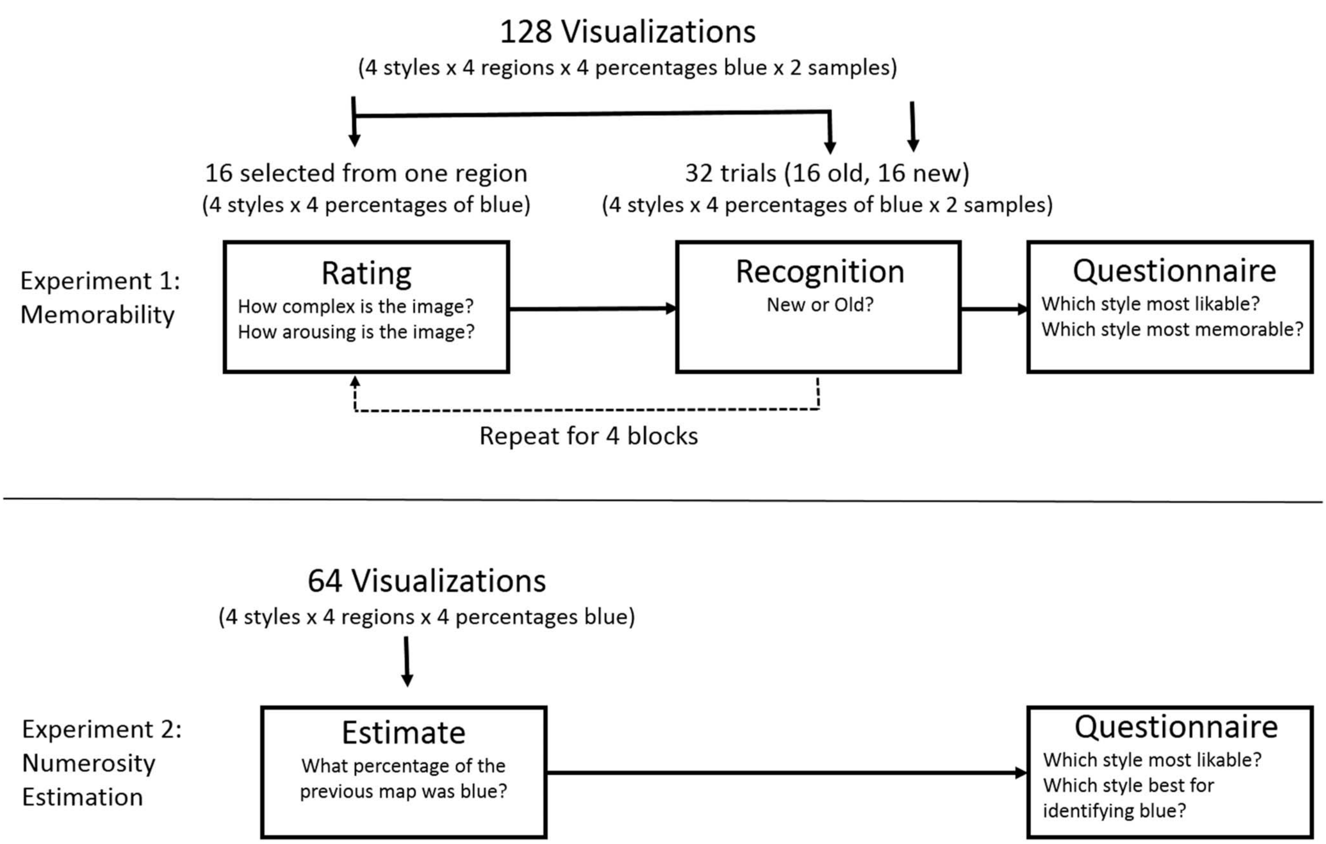

Figure 2. Overview of the experimental design in Experiments 1 and 2.

testing session to ensure that the eye tracker remained calibrated and to answer questions when they arose. The overall experiment was comparable with other studies examining visualization memorability (Borkin et al., 2013, 2016).

Sessions began with instructions being given both verbally and on the screen after which a nine-point eye tracker calibration was completed. Participants finished one practice trial showing a single visualization and a sample question about perceived complexity. Rating trials followed and consisted of a single weather visualization shown in one of the four styles. After viewing a single visualization for five seconds the screen was replaced by one of 16 questions accompanied with a Likert scale that ranged from 1 (*very little*) to 6 (*very much*). Each question asked participants to rate the previous visualization on perceived complexity or arousal (Berlyne, 1971). A total of eight adjectives (Table 1) were used to assess complexity and a separate eight adjectives were used to assess arousal. Representing each dimension with related but different words was done to increase the generality of findings. In addition, half of the adjectives in each set were negatively worded (and scored) to avoid halo effects. The rationale behind varying the words and the directionality of valence was to encourage active exploration of the different visualizations prior to testing their memorability. Adjectives were selected randomly with the constraint that they each occurred an equal number of times per participant. There was no time limit on these responses, though once a response had been entered a fixation cross appeared on the screen for 500ms and the next trial would then begin. Additional 9-point eye tracker calibrations were performed every 10 trials.

After 16 rating trials participants completed a surprise new-old recognition test. At this point participants were informed that they would see 32 visualizations, one at a time, and would be asked to indicate whether they had previously seen the current visualization or not. Compared with the previously rated 16 visualizations, foils were selected to represent the same geographic area, in the same style and depicting similar weather trends. Participants were told that half of the visualizations had been shown in the previous block of trials and half were new. Feedback was not provided and participants were free to view each visualization until a decision was made.

This cycle of 16 rating trials followed by a new-old recognition test involving 32 memory trials was repeated three more times, for a total of 64 rating trials and 128 memory trials. Each rating block depicted images from a single geographic area, with the assignment of area to block being random and without replacement. In doing so, a given block depicted only one geographic area, and within that block only style and weather conditions varied. This allowed us to compare which styles were most closely linked to recognition of specific weather conditions, separately from any influences of geographic area.

After completing the final block of memory trials participants were shown a single visualization depicting the same data in order of glyph, sculptural, containment, and impasto. The styles were respectively labeled 1 to 4, and participants viewed each visualization until pressing the space key. After viewing each style a screen appeared that displayed all four styles together with each having a resolution of 300×300 pixels and accompanied with the previously set 1 to 4 number scheme. Above these visualizations questions appeared one at a time and in the following order: Which of these styles is the most [complex, arousing, likable, and memorable]. Participants responded by typing in a number from 1 to 4.

Data analysis. Signal detection analyses were used to compute d-prime (d') for each visualization style per participant to provide a measure of recognition that is unaffected by tendencies to respond “new” or “old” when participants were unsure. To construct this metric participant hits and false alarms for each visualization style were used. Hits represented trials in which a visualization was correctly identified as previously seen and false alarms as being cases where a visualization was labeled as previously seen when it had not prior been shown. When the proportion of hits or false alarms was on the ceiling or floor, we replaced 1.0 and 0.0 values with 0.99 and 0.001, respectively (Macmillan & Creelman, 2004). A higher value of d' thus represents greater recognition expressed in standard deviation units, such that a d' of 1 means the signal is estimated to be 1 standard deviation unit stronger than noise.

Results

New-old recognition accuracy. The main finding of Experiment 1 was that impasto and glyph visualizations were more memorable than sculptural and containment visualizations. There were no significant differences between impasto and glyphs, or between sculptural and containment. These results are shown in Figure 3 and Table 2.

The main analysis of recognition accuracy was based on the mean d' values of the four styles (see Table 2). A combined analysis of the effects of geographic area and block was not possible because when the data were broken down to that level of detail there were numerous floor and ceiling effects; many cells had a proportion of hits and/or false alarms that were 1.0 or 0, respectively. We therefore used a nonparametric version of recognition accuracy, *A*, to examine geographic area and block (Zhang & Mueller, 2005). These analyses indicated that visualizations of the U.S.A. yielded higher recognition than other geographic areas and that there were no effects of block. All reported conclusions were supported by the analyses described below. In instances of a significant omnibus ANOVA, Fisher’s LSD was used to compare different groups. A Bonferroni correction was applied when multiple comparisons were made in an effort to control alpha. Both the Fisher’s and the Bonferroni p values are reported in these cases.

A repeated measures ANOVA found a significant effect of style (four levels: glyph, sculptural, containment, and impasto) on mean

Table 1
The Adjectives Given to Participants When Rating the Visualization Styles Varied on Two Dimensions (Complexity, Arousal) and Were Equally Often Positive or Negative

Dimension			
Complexity		Arousal	
Positive	Negative	Positive	Negative
Complex	Simple	Arousing	Boring
Intricate	Plain	Stimulating	Stale
Complicated	Easy	Awakening	Sleepy
Elaborate	Basic	Provocative	Dull

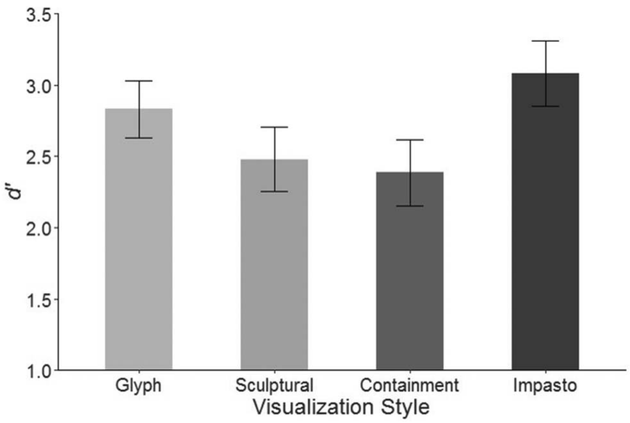

Figure 3. Mean new-old recognition accuracy as indexed by d' in Experiment 1. Error bars represent ± 1 *SEM*.

d' score, $F(3, 87) = 3.85$, $p = .012$, $\eta_p^2 = .12$, with Fisher's LSD showing that the pairing of glyph and impasto yielded significantly higher recognition than the pairing of sculptural and containment, $p = .002$ (Bonferroni $p = .006$). Differences between glyph and impasto however were not found ($p = .281$) nor between sculptural and containment ($p = .686$). An exploratory repeated measures ANOVA found a significant effect of geographic area (four levels: Africa, Asia, South America and the USA) on mean *A* score, ($F(3,87) = 3.79$, $p = .013$, $\eta_p^2 = .12$), for which the USA had highest recognition ($p = .013$). Lastly no effect of block (4 levels: block 1, 2, 3 and 4) was found ($F(3,87) = 2.23$, $p = .091$). We caution that the results concerning continent and block were not hypothesized a priori and should be treated as exploratory.

Perceived functionality & liking. Superior recognition accuracy was found for the styles participants selected as the most liked, but not for those selected as the most memorable. In detail, we computed d' for the visualization style each participant selected as most liked. We then averaged the hits and false alarms of the remaining three styles into a single hit rate and single false alarm rate, from which a d' score was calculated. This averaging of hits and false alarms across styles was done in part to minimize floor and ceiling effects that occur when d' is computed on small samples. We then compared the resulting d' scores for liked versus not liked styles, and similarly memorable versus not memorable styles. Specifically recognition accuracy was higher for visualizations selected as being the most liked, $F(1, 29) = 10.06$, $p = .004$, $\eta_p^2 = .26$, but not those selected as most memorable, $F(1, 29) = 0.56$, $p = .462$. The responses given by participants to questionnaire items is summarized in Table 3.

Visual exploration. Impressionist styles encouraged greater visual exploration than glyphs (see Figure 4). A repeated measures ANOVA revealed a main effect of style on number of fixations, $F(3, 29) = 14.43$, $p < .001$, $\eta_p^2 = .33$, with Fisher's LSD showing that impressionist styles encouraged significantly more fixations than glyphs ($p < .001$).

Complexity & arousal ratings. Complexity and arousal ratings were mostly comparable for the different styles and geographic areas, with the most notable exception being that containment was slightly more complex than the other remaining styles. These results are reported in greater detail in the supplementary material.

Experiment 2: Numerosity Estimation

An effective data visualization reveals data patterns and trends that might otherwise go unnoticed. Participants completed a numerosity estimation task to test which style best depicted the prevalence of a weather condition represented by the color blue.

Method

Participants. Thirty-one different undergraduates (16 males and 15 females, *mean age* = 20.51, *SD* = 2.71) participated, following the same recruitment and consent procedures as in Experiment 1. Eye-tracking data from one participant were excluded because of equipment malfunction. All participants finished the study within one hour.

Stimuli and apparatus. The base visualizations, display equipment, and recording devices were identical to Experiment 1. Visualizations now, however, were further grouped based on what percentage of the visualization was blue: 0–20%, 21–40%, 41–60%, and 61–80%. For each visualization shades of blue were representative of a single weather condition, for instance increasing levels of frost frequency. Each of the four categories of numerosity were displayed an equal number of times for each geographic area and visualization style. Although other studies have used smaller bin categories (e.g., 5%) for visualization distinctions (Cleveland & McGill, 1984), we were limited in that bin categories had to be created from existing weather data sets. We also aimed to maximize perceptible style differences in color proportionality and the four bin ranges seemed to accomplish this well. Since the visualizations use a continuous color scale, a gamut of colors considered to be blue were used. Images were examined

Table 2
Mean Values of Subjective Ratings, d-Prime, and Fixation Count for the Four Styles in Experiment 1 and the Mean Values of Absolute Errors and Fixation Count for Experiment 2

	Experiment 1				Experiment 2	
Style	Arousal	Complexity	*d*-prime	Fixations	Errors	Fixations
Glyph	3.21	3.95	2.83	271.73	.49	279.53
Sculptural	3.21	3.67	2.48	294.50	.40	297.33
Containment	3.42	4.12	2.39	297.77	.41	298.47
Impasto	3.45	3.88	3.08	293.90	.33	299.80

Table 3
The Percentage of Participants in Experiment 1 Who Selected a Visualization Style as Being the Most Complex, Arousing, Likable and Memorable

Style	Complex	Arousing	Likable	Memorable
Glyph	20.00	23.33	13.33	40.00
Sculptural	3.33	20.00	13.33	10.00
Containment	60.00	36.67	26.67	30.00
Impasto	16.67	20.00	46.67	20.00

post-processing to verify the percentage of blue pixels relative to the other color (nonbackground) pixels and adjustments were made as needed. Figure 1 provides for an example in which the correct bin category is 0–20% for all four styles.

Procedure. The procedure was identical to Experiment 1, with the exceptions that on each trial following the five second display of a visualization the screen was replaced with the question "What percentage of the previous map was blue?", and that all visualizations were now presented in an entirely random order, rather than in blocks of trials depicting a single geographic region. Participants responded via keyboard press to select a response option, "1. 0–20%", "2. 21–40%", "3. 41–60%" and "4. 61–80%". No time limit was enforced for response selection.

Instead of relying on a binary outcome (correct or incorrect), each trial was assessed by the degree to which the response given by a participant deviated from the correct answer. An error score was calculated that ranged from 0, indicating the correct answer, to 3, indicating a maximally wrong answer (e.g., estimating 61–80% when in fact the visualization contained a blue percentage between 0 and 20%). At the end of the testing phase, participants completed a questionnaire, analogous to Experiment 1, in which they were asked which of the four visualization styles was the most liked and which style they thought to be most effective for estimating the color blue.

Results

Numerosity accuracy. The main finding was that impasto visualizations resulted in the most accurate estimation of color proportionality. Table 2 and Figure 5 show the mean absolute errors (departures from the correct answer) for each of the four styles. In addition to impasto leading to the most accurate estimates, the three impressionist styles as a group were more accurate than glyphs. These conclusions were supported by the following analyses. As in Experiment 1, Fisher's LSD was used for group comparison following a significant omnibus ANOVA, and a Bonferroni correction was applied in instances of multiple comparisons.

A repeated measures ANOVA on mean absolute errors indicated a significant effect of style ($F(3, 90) = 7.53$, $p < .001$, $\eta_p^2 = .20$). Follow up Fisher's LSD indicated that all three impressionist styles together had greater accuracy than glyphs, $p < .001$ (Bonferroni $p < .001$) and that impasto was superior over the two remaining impressionist styles $p = .021$ (Bonferroni $p = .041$). Though not explicitly hypothesized, a less interesting and expectable main effect of percentage blue (4 levels: 0–20%, 21–40%, 41–60%, 61–80%) was also found, $F(3, 90) = 13.81$, $p < .001$, $\eta_p^2 = .32$, with trials having minimal and maximal blue being easier to correctly estimate than trials with intermediate percentage blue ($p < .001$). An effect of geographic area was also found, $F(3, 90) = 7.60$, $p < .001$, $\eta_p^2 = .20$, with Africa yielding more accurate estimation than the other geographic areas ($p < .001$). We caution that the results concerning continent and percentage blue were not hypothesized a priori and should be treated as exploratory.

Perceived functionality & liking. More accurate numerosity estimation was found for liked styles but not styles thought to be the most effective for task completion. These conclusions were supported by comparing each participant's mean absolute error for the visualization style selected as an answer to a questionnaire item to the average remaining error scores of the nonselected visualization styles. Specifically accuracy was higher for styles selected as most liked, $F(1, 30) = 4.23$, $p = .049$, $\eta_p^2 = .12$, but not for those selected as being most effective for task completion, $F(1, 30) = 1.57$, $p = .220$. These results are similar to Experiment 1 where liking associated with higher task performance but not perceived functionality. The responses given by participants to questionnaire items is summarized in Table 4.

Visual exploration. The main finding was that impressionist styles encouraged greater visual exploration than glyphs (see Fig-

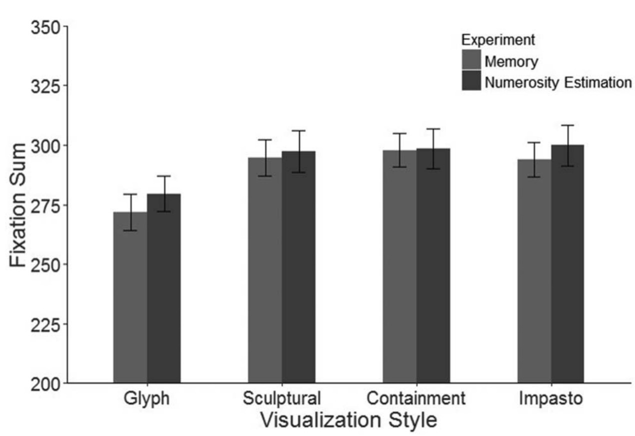

Figure 4. The total number of discrete fixations made when exploring each of the visualization styles in Experiment 1 and 2. Error bars represent ±1 *SEM*.

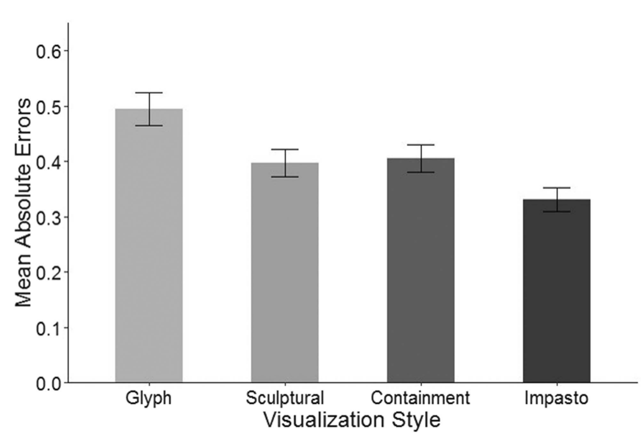

Figure 5. Mean absolute error in estimating the numerosity of the color blue in Experiment 2. Error bars represent ±1 *SEM*.

Table 4
The Percentage of Participants in Experiment 2 Selecting a Visualization Style as Being the Most Likable and Most Effective for Numerosity Estimation

Style	Likable	Effective
Glyph	3.20	6.50
Sculptural	19.40	16.10
Containment	38.70	45.20
Impasto	38.70	32.30

ure 4). A repeated measures ANOVA revealed a main effect of style on number of fixations, $F(3, 29) = 11.12$, $p < .001$, $\eta_p^2 = .28$, with Fisher's LSD showing that impressionist styles encouraged significantly more fixations than glyphs ($p < .001$). These results replicate those obtained in Experiment 1.

General Discussion

The primary questions of this study were whether the painterly techniques of impressionists could produce data visualizations that are memorable, effective in conveying information, are liked, and encourage visual exploration. In addressing these questions we compared three impressionist visualization styles to the frequently used visualization style of glyphs. Experiments were conducted in which participants viewed weather data sets rendered into the four different visualization styles and completed a new-old recognition test (Experiment 1), or a numerosity estimation task (Experiment 2). We anticipated that impressionist visualizations would perform well when compared with glyphs because each employed deliberate techniques to guide viewer attention toward regions of greater data heterogeneity, hence signifying the most distinctive portions of the visualization. At the same time we hypothesized that impasto, which employed multiple techniques known to engage the human visual system, would be more effective than either glyphs or the other impressionist styles. We also examined the relationship between objective performance in each experiment in relation to aesthetic liking and perceived functionality.

The results of the new-old recognition task in Experiment 1 showed that impasto and glyph visualizations were more memorable that the other two impressionist styles: sculptural and containment. A direct comparison of impasto and glyph however showed no difference; neither was there a difference between sculptural versus containment. In Experiment 2 participants estimated the numerosity of the color blue (corresponding to a particular weather condition) in each of the data visualizations and here impressionist visualizations led to greater response accuracy than glyphs, and furthermore impasto outperformed the two remaining impressionist styles.

A secondary hypothesis was whether liked visualizations or those judged to be functional had higher task performance. We explored this via a questionnaire in which participants were asked which style they liked most and which style they thought was most memorable (Experiment 1) or best for numerosity estimation (Experiment 2). We found that liked styles in both Experiment 1 and Experiment 2 were indeed those that resulted in better task performance, whereas the styles selected as being the most effective for task completion did not. Relevant here is research by Da Silva, Crilly, and Hekkert (2017) who identified the principle of *maximum effect for minimum means*. In short, whether an object is appreciated or not will depend on how effective and functional it is relative to other objects. Thus when participants were asked which style they most liked their answer may have been influenced by their recent experience of successful task completion. As might be expected from this principle, styles that performed well in both experiments also tended to be those that were most liked. This may suggest that participants did not strictly rely on aesthetic judgment when selecting the style they liked, but rather may have drawn from an intuitive sense of task performance not captured by explicitly perceived functionality.

Although both Experiments 1 and 2 found a link between liking and effectiveness, consistent with previous reports (Tractinsky et al., 2000; Norman, 2005), we caution that the directionality of this effect cannot be established from these data (Tuch, Roth, Hornbæ, Opwis, & Bargas-Avila, 2012). A style may be liked because it has been associated with a successful task experience (Makin, Wilton, Pecchinenda, & Bertamini, 2012) or because selective attention has been paid to it during target identification (DiPaola et al., 2010). Indeed, there is ample evidence to support the hypothesis that attention and liking are linked in a bidirectional way (see review by Brennan & Enns, 2014). Simply paying more attention to an event seems to increase its subjective value and similarly, increases in the subjective value of a target leads to greater selective attention to it. Future studies should be able to examine this bidirectionality by asking participants to rate stimuli both before and after task completion. The critical question would be whether aesthetic ratings change as a function of task performance. An even more direct experimental approach would be to see whether aesthetic liking could be manipulated via false feedback on task performance.

A third question addressed in this study was whether impressionist visualizations, by virtue of borrowing the painterly techniques of impressionist masters, would elicit greater visual exploration. In both experiments glyphs tended to be liked less than the other visualization styles, and glyphs also tended to elicit the least visual exploration. This finding is consistent with previous research indicating that liking also associates with greater visual exploration (Maughan et al., 2007).

Limitations

Some of the findings reported here were not central to our theoretical focus but are worth considering briefly. We did not for instance anticipate that visualizations of the United States would yield higher new-old recognition overall (Experiment 1) or that the African continent would have better numerosity estimation (Experiment 2). As such, we can only speculate on these effects. Perhaps the United States was more easily recognized given that the participants were North American university students and as such have greater familiarity (perhaps in shape, weather trends, or other) with images of the United States (Hegarty, Canham, & Fabrikant, 2010). By the same reasoning, the question of why Africa allowed for more accurate numerosity estimation in Experiment 2 is puzzling. We note that Africa also yielded the highest arousal ratings in Experiment 1 (see supplementary material; Figure S1) and perhaps this intrigue associated with closer image inspection and accurate estimation. Although we have provided speculative interpretations of the effects of geographic area, we remind readers that no specific hypotheses were proposed a priori and so these findings are to be treated as exploratory.

In Experiment 1 participants rated each visualization style on complexity and arousal with the most notable difference being that containment was in both the study phase and the poststudy questionnaire rated as most complex. We speculate that because individual brushstrokes in this style were so heavily outlined and emphasized the resulting image became overly detailed. Such an interpretation would seem to agree with advice of using as little ink as possible to depict data trends, which in turn this style may have violated through such accentuated outlines (Tufte, 1983). One speculative interpretation is that intermediate complexity is ideal as found in the two more memorable styles, glyphs and impasto (Berlyne, 1971).

In Experiment 2 participants identified a weather property by estimating color proportionality. Building on our design participants may explicitly be informed what data property is being mapped onto color. This would allow for a more direct examination if weather properties, and not just color features, were being more accurately estimated. Experiment 2 was also limited in that only a single weather property was tested. To more closely approximate real-world conditions in which multiple weather conditions are often of interest, future studies should request participants to estimate, for example, color proportionality and orientation (of brushstrokes or glyphs) thus allowing multiple weather properties to be judged concurrently.

Because participants were asked which style they liked most, rather than asking for their subjective ratings on every style, we could not thoroughly compare the results of both experiments. This also meant we could also not compare the gradient of liking across all four styles relative to performance. The results showed, nonetheless, that in both Experiments 1 and 2, impasto was chosen by a majority of the participants as the most liked styles and glyphs were clearly among the least liked. Using an interval rating scale in future work would though better help differentiate and rank the styles. Better understanding may also be achieved if participants were asked to identify which features of a style they thought lead to increased liking.

Why Impasto?

There may be many reasons why impasto performed especially well. First, this style is unique in using selective variation to highlight image regions that are locally distinct from their surroundings (DiPaola et al., 2010; Ramachandran & Hirstein, 1999), and within these regions, this style adds further variation via within-stroke differences. Portions of the visualization are especially emphasized through the technique from which the style is named, as described by van Gogh

> Good painting does not depend upon using much color, but to paint a ground with force, or to keep a sky clear, one must sometimes not spare a tube. Sometimes the subject requires a delicate painting; at times the nature of things themselves requires a thick painting.

The resulting visualizations are both distinct and have an almost decorative appearance from which they likely benefited (Bateman et al., 2010; Borgo et al., 2012).

Second, the brushstrokes of this style have the greatest amount of round curvature by angle. In contrast, the style containment had the harshest angles often appearing to have sharp and square edges, with sculptural brushstrokes falling somewhere in-between. Glyphs too were composed of rectilinear like shapes and often appeared sharp. Such curvature differences likely play a role in the aesthetic appeal of impasto, as smoothness and rounder contour have been found to be visually pleasing for both abstract and real-world objects (Bar & Neta, 2006; Bertamini, Palumbo, Gheorghes, & Galatsidas, 2016).

A final consideration arises from Wood and colleagues (2012) who noted that visualizations may differ in their perceived purpose. A visualization very clearly made by a computer algorithm may signify to the viewer less purpose or significance; that the visualization was not created with a specific intention but rather is just the end product of a robotic process (Kirk, Skov, Hulme, Christensen, & Zeki, 2009). Alternatively, a visualization that appears handcrafted or uniquely designed may imply manual effort and that the visualization was built with a specific goal or purpose to achieve. Future studies may consider asking participants the degree to which they believe a particular visualization had been personalized and made with effort, and whether this, in turn, influences aesthetic ratings or motivation to perform well on the current task.

A final word should be made about our use of impressionist painting techniques for a purpose they were never intended for, namely, data visualization. It is important to remind readers that impressionist techniques were originally developed to depict pastoral scenes, portraits, and landscapes, not abstract data patterns as in the present study. Nonetheless, because these techniques were historically effective in the domain of realism, we think it is worth asking whether the same techniques might also be effective in the more abstract domain of spatially arrayed data. As our results suggest, the techniques of masterful artists have broad appeal and functional use for something as alien as depicting weather trends. Although this better understanding may have been first achieved in a trial and error fashion by master artists in the impressionist school, we think it is now worth exploring them more systematically to harness their power in a scientific context.

References

Bar, M., & Neta, M. (2006). Humans prefer curved visual objects. *Psychological Science, 17,* 645–648. http://dx.doi.org/10.1111/j.1467-9280.2006.01759.x

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

Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York, NY: McGraw-Hill.

Bertamini, M., Palumbo, L., Gheorghes, T. N., & Galatsidas, M. (2016). Do observers like curvature or do they dislike angularity? *British Journal of Psychology, 107,* 154–178.

Borgo, R., Abdul-Rahman, A., Mohamed, F., Grant, P. W., Reppa, I., Floridi, L., & Chen, M. (2012). An empirical study on using visual embellishments in visualization. *IEEE Transactions on Visualization and Computer Graphics, 18,* 2759–2768. http://dx.doi.org/10.1109/TVCG.2012.197

Borgo, R., Kehrer, J., Chung, D. H., Maguire, E., Laramee, R. S., Hauser, H., . . . Chen, M. (2013). Glyph-based visualization: Foundations, design guidelines, techniques and applications. *Eurographics State of the Art Reports,* 39–63.

Borkin, M. A., Bylinskii, Z., Kim, N. W., Bainbridge, C. M., Yeh, C. S., Borkin, D., . . . Oliva, A. (2016). Beyond Memorability: Visualization Recognition and Recall. *IEEE Transactions on Visualization and Computer Graphics, 22,* 519–528. http://dx.doi.org/10.1109/TVCG.2015.2467732

Borkin, M. A., Vo, A. A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., & Pfister, H. (2013). What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics, 19,* 2306–2315. http://dx.doi.org/10.1109/TVCG.2013.234

Brennan, A. A., & Enns, J. T. (2014). Visual attention and emotion: Studying influence on a two-way street. In R. R. Hoffman, P. A. Hancock, M. Scerbo, R. Parasuraman, & J. L. Szalma (Eds.), *The Cambridge handbook of applied perception research* (pp. 175–198). New York, NY: Cambridge University Press.

Brown, R. (1978). Impressionist technique: Pissarro's optical mixture. In B. E. White (Ed.), *Impressionism in perspective* (pp. 114–121). Englewood Cliffs, NJ: Prentice-Hall.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think*. Burlington, MA: Morgan Kaufmann.

Cavanagh, P. (2005). The artist as neuroscientist. *Nature, 434,* 301–307. http://dx.doi.org/10.1038/434301a

Chen, C. (2005). Top 10 unsolved information visualization problems. *IEEE Computer Graphics and Applications, 25,* 12–16. http://dx.doi.org/10.1109/MCG.2005.91

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association, 79,* 531–554. http://dx.doi.org/10.1080/01621459.1984.10478080

Cutting, J. E. (2006). *Impressionism and its canon*. Lanham, MD: University Press of America.

Da Silva, O., Crilly, N., & Hekkert, P. (2017). Beauty in efficiency: An experimental enquiry Into the principle of maximum effect for minimum means. *Empirical Studies of the Arts, 35,* 93–120. http://dx.doi.org/10.1177/0276237416638488

DiPaola, S., Riebe, C., & Enns, J. T. (2010). Rembrandt's textural agency: A shared perspective in visual art and science. *Leonardo, 43,* 145–151. http://dx.doi.org/10.1162/leon.2010.43.2.145

DiPaola, S., Riebe, C., & Enns, J. T. (2013). Following the masters: Portrait viewing and appreciation is guided by selective detail. *Perception, 42,* 608–630. http://dx.doi.org/10.1068/p7463

Fabrikant, S. I., Christophe, S., Papastefanou, G., & Maggi, S. (2012, September). *Emotional response to map design aesthetics*. 7th International Conference on Geographical Information Science, Columbus, OH.

Fabrikant, S. I., Hespanha, S. R., & Hegarty, M. (2010). Cognitively inspired and perceptually salient graphic displays for efficient spatial inference making. *Annals of the Association of American Geographers, 100,* 13–29. http://dx.doi.org/10.1080/00045600903362378

Harrison, L., Reinecke, K., & Chang, R. (2015). Infographic aesthetics: Designing for the first impression. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1187–1190). ACM.

Healey, C. G., & Enns, J. T. (2012). Attention and visual memory in visualization and computer graphics. *IEEE Transactions on Visualization and Computer Graphics, 18,* 1170–1188. http://dx.doi.org/10.1109/TVCG.2011.127

Healey, C. G., Tateosian, L., Enns, J. T., & Remple, M. I. (2004). Perceptually based brush strokes for nonphotorealistic visualization. *ACM Transactions on Graphics, 23,* 64–96. http://dx.doi.org/10.1145/966131.966135

Hegarty, M., Canham, M. S., & Fabrikant, S. I. (2010). Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 37–53. http://dx.doi.org/10.1037/a0017683

Hogarth, W. (1753). *The analysis of beauty written with a view of fixing the fluctuating ideas of taste*. London, UK: Reeves.

Intergovernmental Panel on Climate Change. (2014). *Climatic research unit global climate dataset data distribution center*. Retrieved from http://www.ipcc-data.org/observ/clim/cru_climatologies.html

Joshi, A., Caban, J., Rheingans, P., & Sparling, L. (2009). Case study on visualizing hurricanes using illustration-inspired techniques. *IEEE Transactions on Visualization and Computer Graphics, 15,* 709–718. http://dx.doi.org/10.1109/TVCG.2008.105

Kirby, J., Stonor, K., Roy, A., Burnstock, A., Grout, R., & White, R. (2003). Seurat's painting practice: Theory, development and technology. *The National Gallery Technical Bulletin, 24,* 4–37.

Kirk, U., Skov, M., Hulme, O., Christensen, M. S., & Zeki, S. (2009). Modulation of aesthetic value by semantic context: An fMRI study. *NeuroImage, 44,* 1125–1132. http://dx.doi.org/10.1016/j.neuroimage.2008.10.009

Koenderink, J., van Doorn, A., & Wagemans, J. (2012). Picasso in the mind's eye of the beholder: Three-dimensional filling-in of ambiguous line drawings. *Cognition, 125,* 394–412. http://dx.doi.org/10.1016/j.cognition.2012.07.019

Latto, R., Brain, D., & Kelly, B. (2000). An oblique effect in aesthetics: Homage to Mondrian (1872–1944). *Perception, 29,* 981–987. http://dx.doi.org/10.1068/p2352

Latto, R., & Russell-Duff, K. (2002). An oblique effect in the selection of line orientation by twentieth century painters. *Empirical Studies of the Arts, 20,* 49–60. http://dx.doi.org/10.2190/3VEY-RC3B-9GM7-KGDY

Lau, A., & Moere, A. (2007). Towards a model of information aesthetics in information visualization. In *Information Visualization, 2007. 11th International Conference* (pp. 87–92). Piscataway, NJ: IEEE.

Lee, M. D., Butavicius, M. A., & Reilly, R. E. (2003). Visualizations of binary data: A comparative evaluation. *International Journal of Human-Computer Studies, 59,* 569–602. http://dx.doi.org/10.1016/S1071-5819(03)00082-X

Lu, A., Morris, C. J., Ebert, D. S., Rheingans, P., & Hansen, C. (2002). Non-photorealistic volume rendering using stippling techniques. In *Visualization, 2002* (pp. 211–218). Piscataway, NJ: IEEE.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. London, UK: Psychology Press.

Makin, A. D., Wilton, M. M., Pecchinenda, A., & Bertamini, M. (2012). Symmetry perception and affective responses: A combined EEG/EMG study. *Neuropsychologia, 50,* 3250–3261. http://dx.doi.org/10.1016/j.neuropsychologia.2012.10.003

Maughan, L., Gutnikov, S., & Stevens, R. (2007). Like more, look more. Look more, like more: The evidence from eye-tracking. *Journal of Brand Management, 14,* 335–342. http://dx.doi.org/10.1057/palgrave.bm.2550074

Müller, H., Reihs, R., Zatloukal, K., & Holzinger, A. (2014). Analysis of biomedical data with multilevel glyphs. *BMC Bioinformatics, 15,* S5. http://dx.doi.org/10.1186/1471-2105-15-S6-S5

Norman, D. A. (2005). *Emotional design: Why we love (or hate) everyday things*. New York, NY: Basic books.

Pezdek, K., & Chen, H. C. (1982). Development differences in the role of detail in picture recognition memory. *Journal of Experimental Child Psychology, 33,* 207–215. http://dx.doi.org/10.1016/0022-0965(82)90016-9

Pezdek, K., Maki, R., Valencia-Laver, D., Whetstone, T., Stoeckert, J., & Dougherty, T. (1988). Picture memory: Recognizing added and deleted details. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 468–476. http://dx.doi.org/10.1037/0278-7393.14.3.468

Pileggi, H., Stolper, C. D., Boyle, J. M., & Stasko, J. T. (2012). Snapshot: Visualization to propel ice hockey analytics. *IEEE Transactions on*

Visualization and Computer Graphics, 18, 2819–2828. http://dx.doi.org/10.1109/TVCG.2012.263

Ramachandran, V. S., & Hirstein, W. (1999). The science of art: A neurological theory of aesthetic experience. *Journal of Consciousness Studies, 6,* 15–51.

Ropinski, T., Oeltze, S., & Preim, B. (2011). Survey of glyph-based visualization techniques for spatial multivariate medical data. *Computers & Graphics, 35,* 392–401. http://dx.doi.org/10.1016/j.cag.2011.01.011

Schapiro, M. (1997). *Impressionism: Reflections and perceptions*. New York, NY: Braziller.

Tateosian, L. G. (2006). *Investigating aesthetic visualizations* (Doctoral dissertation). Retrieved from http://www.lib.ncsu.edu/resolver/1840.16/3832

Tateosian, L. G., Healey, C. G., & Enns, J. T. (2007). Engaging viewers through nonphotorealistic visualizations. In *Proceedings Fifth International Symposium on Non-Photorealistic Animation and Rendering* (pp. 93–102). New York, NY: ACM.

Tractinsky, N., Katz, A. S., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers, 13,* 127–145. http://dx.doi.org/10.1016/S0953-5438(00)00031-X

Tuch, A. N., Roth, S. P., Hornbæ, K. K., Opwis, K., & Bargas-Avila, J. A. (2012). Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI. *Computers in Human Behavior, 28,* 1596–1607. http://dx.doi.org/10.1016/j.chb.2012.03.024

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

van Gogh, V. (1937). (I. Stone, Ed.). *Dear Theo: The autobiography*. Edinburgh, UK: Constable.

Ward, M. O. (2008). Multivariate data glyphs: Principles and practice. In C. Chen, W. Haerdle, & A. Unwin (Eds.), *Handbook of data visualization* (pp. 179–198). Berlin, Germany: Springer Berlin Heidelberg. http://dx.doi.org/10.1007/978-3-540-33037-0_8

Ware, C., & Plumlee, M. D. (2013). Designing a better weather display. *Information Visualization, 12,* 221–239. http://dx.doi.org/10.1177/1473871612465214

Wood, J., Isenberg, P., Isenberg, T., Dykes, J., Boukhelifa, N., & Slingsby, A. (2012). Sketchy rendering for information visualization. *IEEE Transactions on Visualization and Computer Graphics, 18,* 2749–2758. http://dx.doi.org/10.1109/TVCG.2012.262

Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika, 70,* 203–212. http://dx.doi.org/10.1007/s11336-003-1119-8

Received March 21, 2017
Revision received January 10, 2018
Accepted February 12, 2018 ■