

Comparing Dot and Landscape Spatializations for Visual Memory Differences

Melanie Tory, Colin Swindells, and Rebecca Dreezer

Abstract—Spatialization displays use a geographic metaphor to arrange non-spatial data. For example, spatializations are commonly applied to document collections so that document themes appear as geographic features such as hills. Many common spatialization interfaces use a 3-D landscape metaphor to present data. However, it is not clear whether 3-D spatializations afford improved speed and accuracy for user tasks compared to similar 2-D spatializations. We describe a user study comparing users' ability to remember dot displays, 2-D landscapes, and 3-D landscapes for two different data densities (500 vs. 1000 points). Participants' visual memory was statistically more accurate when viewing dot displays and 3-D landscapes compared to 2-D landscapes. Furthermore, accuracy remembering a spatialization was significantly better overall for denser spatializations. These results are of benefit to visualization designers who are contemplating the best ways to present data using spatialization techniques.

Index Terms—Information interfaces and presentation, screen design, evaluation / methodology, user / machine systems, software psychology, landscape visualization.

1 INTRODUCTION

This work empirically demonstrates that two common abstract data visualizations, landscapes and contour plots, are not more memorable compared to dot displays. Three-dimensional (3-D) 'landscape' displays and two-dimensional (2-D) contour plots or 'heat maps' are popular ways to visualize some types of non-spatial data. For example, the Inspire system arranges text articles in a terrain-like display to illustrate themes [4, 31]. There has been some uncertainty in the field about when these types of visualizations improve user performance compared to simpler dot-based displays, with prior experiments showing mixed results.

One challenge with landscape and contour displays is that the user must sometimes remember the overall shape of a landscape in order to compare it to another version. For instance, Nowell et al. [21] report that when landscape views are shown for different time slices, users must retain one landscape in memory in order to compare it to the next slice and understand how the landscape has changed. Although some techniques have been developed to reduce memory load when comparing two landscapes [21], comparing more than two is more challenging and the memory problem returns. Our experiment examined how simple display design decisions impact a user's ability to remember a layout. In particular, we suspected that the abstraction provided by landscapes or contour plots would improve memorability compared to direct point displays. This conjecture turns out to be untrue, as we show later.

Dot, landscape, and contour displays fall into a class of displays called spatializations [10], which show abstract, non-spatial data using a geographic metaphor. Document collections are the most common use of spatializations. However, spatialization can also be applied to multidimensional data such as a table of automobile statistics. The spatial arrangement of items in a spatialization is

typically created through a dimensionality reduction technique such as multi-dimensional scaling (MDS) or principle component analysis (PCA). These algorithms reduce the high dimensional space to a lower dimensionality that can be displayed on a computer screen, usually 2-D. Such layouts allow the user to infer the similarity of items by observing their spatial distance (the distance-similarity metaphor). Research suggests that spatializations promote understanding of high dimensional relationships, by enabling users to easily see similarities, clusters, and outliers [1, 2, 13].

Once the 2-D layout of points has been determined, one needs to decide how to visually represent the spatialization. A common approach is to create a 3-D information landscape by fitting a surface to the points. With this approach, coloured 'hills' represent peaks of the visualized data. Various properties of the data can be mapped to the height dimension. For instance, in a collection of documents, the height could be used to represent the local point density (i.e. the number of documents), or in an automobile dataset, the height may represent a data dimension such as fuel economy. Other possible representations are a 2-D landscape (where 'height' is represented by colour and/or contours similar to a topographic map) and coloured dots (where no surface is fitted to the points, and data points are simply coloured according to the property of interest).

We present an experiment comparing these different visual representations of spatializations. We categorize spatializations into two groups based on the graphical mark used to represent data (see Figure 1):

- *Dots*: Spatializations that show only dots (points).
- *Information Landscapes*: Spatializations where a surface has been fitted to the set of underlying points. Dots may be shown on the surface, and the surface may be flat (2-D) or may vary in height (3-D). We refer to these simply as landscapes.

Our experiment specifically compares 2-D and 3-D landscapes to dot-based displays for a memory task. Because landscapes provide an abstracted view, we expected them to be easier to remember than dot-based displays. A representation that is easier to remember may also serve as a better 'mental map' of the overall structure. An effective mental map could be particularly important in zooming interfaces, where the user cannot always see an overview. The importance of mental maps in visualization has been known for some time (e.g., Misue et al. [18]). More importantly, users often need to remember spatializations in order to make mental comparisons and understand high-level differences among many different views.

-
- *Melanie Tory is with the University of Victoria, E-Mail: mtory@cs.uvic.ca.*
 - *Colin Swindells is with the University of Victoria, E-Mail: colinedw@cs.uvic.ca.*
 - *Rebecca Dreezer is with McMaster University, E-Mail: dreezerj@muss.cis.mcmaster.ca.*

Manuscript received 31 March 2009; accepted 27 July 2009; posted online 11 October 2009; mailed on 5 October 2009.

For information on obtaining reprints of this article, please send email to: tvcg@computer.org .

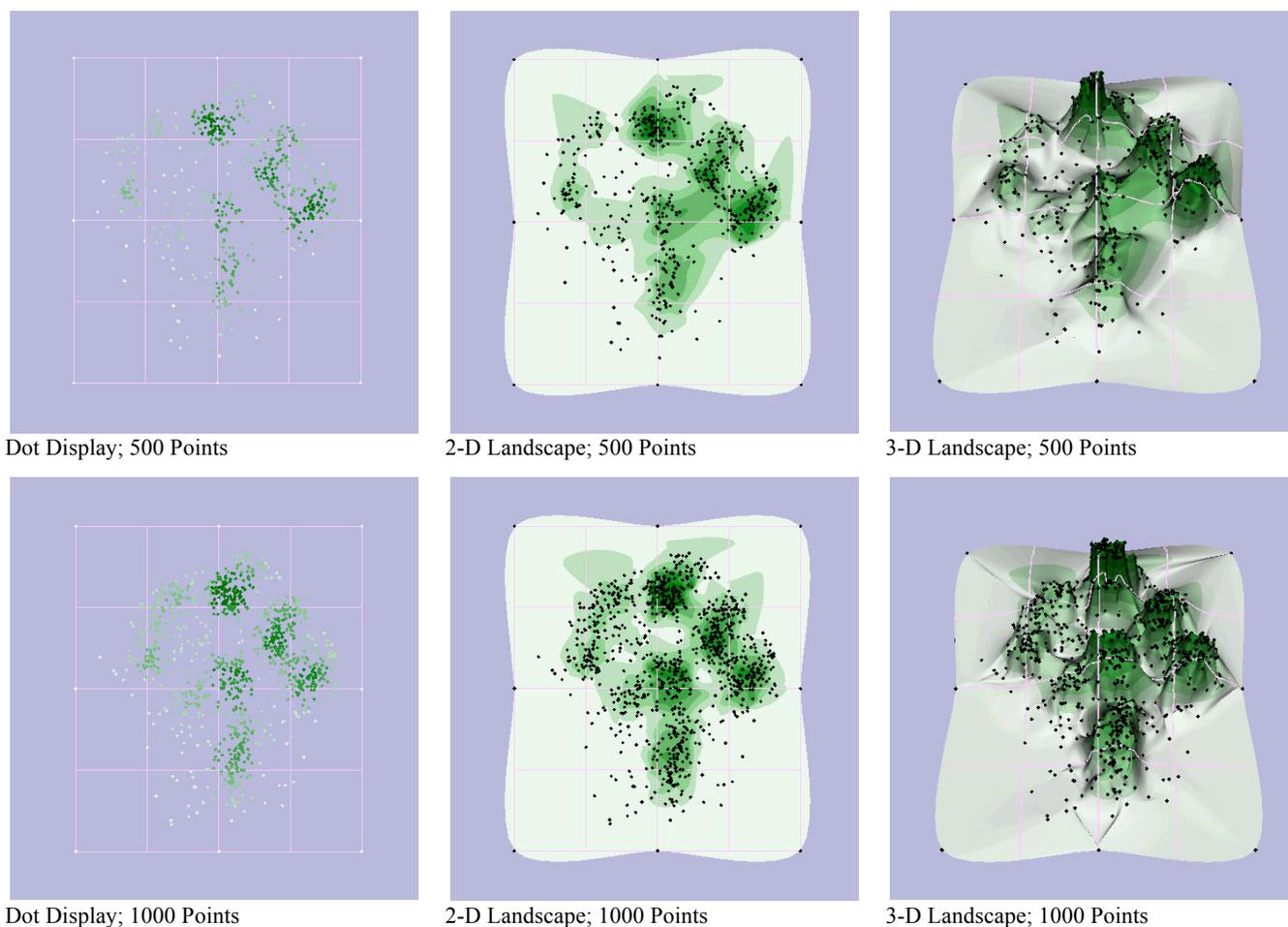


Figure 1: Six display types of the user study. All show the same data. Colour saturation represents point density, ranging from white (low) to dark green (high).

1.1 Common Assertions about Landscapes

Despite the known challenges associated with 3-D displays, such as navigation and occlusion, many researchers have proposed using landscapes (e.g., [1, 4, 10, 14, 31]). Proponents of landscapes suggest several reasons why 2-D and 3-D landscapes may be beneficial. Using a landscape metaphor may facilitate pattern recognition and spatial reasoning [26]. Information landscapes may avoid problems with some dimensions obfuscating others, may simplify the amount of data to be presented, and may display data in a way that is optimal for information processing [4]. Several authors suggest that the landscape metaphor is easily understood by most users and facilitates hierarchical clustering of data items [1, 13, 14, 15]. For 3-D spatializations, the landscape surface may also provide a constant reference to reduce disorientation when navigating in 3-D space [4] and to aid depth perception [20]. Of particular interest to our study, some researchers postulate that landscapes' abstraction and spatial metaphor make the layout easier to remember, perhaps because it is more similar to our everyday world [4]. Whether or not this assertion is true is unclear based on prior research. In general, most of these assertions about the benefits of information landscapes have not been tested empirically.

1.2 Empirical Knowledge about Graphical Encoding and Visual Memory

Redundantly encoding data using two or more retinal variables has been shown to improve perceptual salience and task performance for some retinal variables and tasks [3, 22], and is a well-accepted

design principle. However, it remains unclear whether 3-D landscapes that redundantly encode data using height with colour have similar benefits. 3-D displays often suffer from occlusion and clutter, and can be difficult to interact with. For example, Cockburn and McKenzie [7] found that 3-D worlds were more difficult to perceive and analyze. For these reasons, redundantly encoding information using height plus colour may actually be detrimental. In our prior work [28], we found that this was indeed the case for a numerosity task with information landscapes. However, we suspected that redundant encoding might be helpful for other mental operations, particularly visual memory.

We focus on the mental operation of remembering a spatialization's overall shape. Previous research has compared 2-D and 3-D interfaces for memory operations, but for very different applications and interface designs. Early studies on document storage and retrieval [23] and memory for objects in a hierarchy [27] found better performance with 3-D displays compared to 2-D. However, more carefully controlled replicates of these experiments found that dimensionality made no significant difference [5, 6, 7]. These later results suggest that 3-D displays do not improve spatial memory; however, it is not clear whether or not these results extend to 2-D and 3-D spatializations. Spatializations differ from the 3-D interfaces previously studied because the third dimension is used to encode data, usually redundantly with other cues such as colour. This redundant encoding may provide an additional aid to memory beyond the basic 3-D metaphor.

Other research has shown that users' ability to recognize objects decreases when the scene undergoes transformation such as scaling, rotation, or fisheye distortion [9, 17, 25]. Because rotatable 3-D

views would likely suffer from these transformation effects, we kept all views static in our experiment.

2 RELATED SPATIALIZATION RESEARCH

Dot spatializations have been well studied, and research has found 2-D dot-based displays to be effective for several tasks. Montello et al. [19] demonstrated that people naturally equate distance with similarity, verifying the distance-similarity metaphor. However, they also found that visual illusions and clusters could override the distance-similarity metaphor. Hornbæk and Frøkjær [15] reported that information retrieval using a 2-D dot display was not more effective than using text summaries, but subjects preferred the spatialization. By contrast, 3-D dot spatializations have been shown to have serious usability problems due to occlusion [4], scene complexity [4], depth ambiguity [20], and difficulty of 3-D navigation [20]. Direct comparisons of 2-D and 3-D dot spatializations [10, 30], have consistently reported better user performance with 2-D dots than with 3-D dots. Westerman and Cribbin [29] also found that better layout algorithms improved performance. For this reason, we do not consider 3-D dot displays in our study.

For landscape spatializations, experiments have reported mixed results. Fabrikant [10] demonstrated that people could intuitively understand the distance-similarity metaphor, landscape representations of non-spatial data, and the relationship between 3-D landscapes and the underlying data points. She found that 2-D landscapes were usually faster than 3-D landscapes for simple distance and density judgment tasks, but that accuracy of some tasks was higher with 3-D than 2-D. She also found that both dots and landscapes were quite effective for distance judgment, but dots were less effective than landscapes for density judgment. In our prior work [28], we found that dots outperformed landscapes for a search and dot estimation task, and that 2-D landscapes outperformed 3-D landscapes. These mixed results suggest that different representations may be suitable for different mental operations, as one might expect. Our experiment extends empirical knowledge in this area by examining a new mental operation: visual memory.

Other variations of spatialized displays have also been considered empirically. Fabrikant et al. investigated perceptual issues in spatializations with graphical links between objects [12] and in spatializations similar to choropleth maps [11]. Cribbin and Chen [8] demonstrated that visually connecting similar nodes in a dot spatialization improved performance at some tasks. We did not consider these more specialized types of spatializations in our study.

3 OUR EXPERIMENT

We designed an experiment to compare memorability of 2-D landscapes, 3-D landscapes, and dot-based spatializations. Our long-term research objective is to develop guidelines for spatialization design by examining how design characteristics influence suitability for different low-level mental operations (e.g. mental transformation or estimation) and higher-level visualization tasks. In this study, we compared three spatialization designs for their ability to support mental operations involving visual memory. Visual memory plays an important role in higher-level tasks where spatializations need to be mentally compared.

We chose the representations shown in Figure 1 because they were the most likely to be effective data display methods. We did not consider dots with no colour mapping because they are unable to encode data outside of spatial position. We also did not consider 3-D dots or 2-D landscapes without colour because they have previously been shown to be ineffective for many tasks [10]. We chose to use colour and height to represent *point density* (as opposed to data values from one dimension). We made this decision because point density is commonly represented this way in spatialization interfaces and because we felt it was the more conservative choice. Point density is redundantly encoded in all of our displays via point positions. If any difference in memorability can be found in this

redundant encoding situation, we would expect the difference to be even greater when the redundant encoding is not present.

In addition to our main factor of display type, we also compared two data densities, to determine whether the abstraction provided by landscapes provided greater benefit as the number of points increased. Based on pilot tests, we chose 1000 points as the larger size; with more points the landscapes were too occluded. We then chose half this number to be the smaller point size. Although these numbers are small compared to many datasets in use today, we believed that in practice larger datasets would be filtered to reduce the visible points, or have the points hidden altogether such that only an abstracted surface was shown.

Our displays were designed to include only the most salient visual features currently found in spatialization displays. This minimalist design enabled us to carefully control differences between conditions. Figure 1 illustrates the six displays compared in our study.

We designed our study to answer the following questions:

- Which are easier to remember: landscapes or dots?
- In landscapes, does redundantly encoding data using colour and height improve memory compared to color alone?
- How is memory of spatializations influenced by the point density in the display?

4 EXPERIMENT DESIGN

This section describes our experimental design. The main factor of display type (Display) had 3 levels: 3-D landscape, 2-D landscape, and dots. The secondary factor of data density (Density) had two levels: 500 points and 1000 points. Examples of each condition are shown in Figure 1. We used a within-subjects design, and order of the six conditions was randomized. Measures were response time and accuracy.

4.1 Participants

Thirty paid participants (11 female, 19 male) were individually tested in the study that took approximately 30 minutes to complete. Their ages ranged from 18-25 years ($M = 21$, $SD = 2.03$). Each participant had normal or corrected to normal vision, and was required to pass an Ishihara color-blindness test [16] to qualify as a participant. Participants were financially compensated for their time.

4.2 Task

For each condition, participants performed a two-phase memory task. In the first phase (learning), the participant viewed eight images successively and was asked to remember them. In the second phase (testing), the participant viewed eight images successively, and was asked to indicate whether or not each image was present in the learning set. Half of the testing phase images were present in the training set and half were new. This same experimental method was used in previous research on memory of node-link diagrams [17].

Learning phase images were each displayed for 12 seconds and followed by a 2.5-second blank screen before the next image appeared, as in previous work [17]. Participants were warned that they would need to recognize these images in the testing phase. Testing phase images appeared until the user pressed one of two keyboard keys to indicate whether or not the image had been seen before. Response times and correctness were recorded.

4.3 Stimuli

For each condition, we created twelve images, each based on a unique multidimensional dataset. We used real rather than synthetic data to ensure the spatializations were realistic. Participants were not told anything about the nature of the data. To ensure consistency, we used the same data sets for all conditions.

Two sizes of each dataset were created, one with 1000 points and one with 500 points. Each 1000-point set consisted of randomly

chosen points from the real data set, and each 500-point dataset was a random subset of the 1000-point dataset. The upper data set size was chosen to be as large as possible without causing substantial occlusion of the landscape surface or other points. These data sets were then laid out in 2-D space using the MDSteer system for multidimensional scaling [30]. Only one layout of each data set was used, to ensure that each display condition showed the same layout. Care was taken to ensure that the twelve images for each condition could be easily distinguished from one another (i.e., they were not too similar).

Point density was the variable we chose to visually represent, as described in section 3. A local density value was calculated for each point and displayed graphically using:

- Dot colour (for dot conditions),
- Surface colour (for all landscape conditions), and / or
- Surface height (for 3-D landscape conditions).

Point density was calculated by (1) dividing the layout canvas into square grid cells, (2) totalling the number of points in each grid cell, and then (3) smoothing the data by assigning each cell the average of itself and its closest neighbours. This result was normalized according to the total number of points in the display. Low density values appeared as valleys in the 3-D displays and high values appeared as peaks. A seven-level saturation colour scale represented point density, ranging from white (low density) to dark green (high density).

For dot-based displays, points were directly displayed and were assigned a colour based on their density value. For the other displays, a graphic surface was created. For 3-D displays, points were first moved to a height representing their density value. Points were then triangulated to create a surface and the surface was smoothed to make it appear like continuous terrain. Contours were extracted at interpolated boundaries between density value ranges, and shown using colour bands. Points were coloured black in the landscape displays. Point colour was chosen to maximize visibility. These color-coding schemes for Dot-based, 2-D landscape, and 3-D landscape spatializations were chosen to most closely match likely practical implementations of each spatialization.

Dot-based displays and 2-D landscapes were rendered from a bird's-eye viewpoint and 3-D landscapes were rendered from an oblique viewpoint, as shown in Figure 1. The oblique viewpoint was chosen to reveal the 3-D nature of the landscape while avoiding excessive occlusion of peaks. All images were static.

The eight learning phase images were randomly selected from the twelve possible images. The testing phase showed the remaining four images, plus four randomly selected images from the learning phase. Images were presented in random order during both phases.

4.4 Apparatus

Visual stimuli were created using custom software that was written in Java using the Visualization Toolkit [24]. Stimuli were saved as static images and presented to participants using DirectRT software. Experiments were run using an AMD Athlon 64 bit dual core PC running at 2.2 GHz, with 2 GB of RAM and Windows Vista. The display was a 21" LCD at 1600 x 1200 resolution with a 32 bit sRGB colour mode. Participants interacted with the software using a standard keyboard at a desk in a dimmed experiment room.

4.5 Procedure

Participants were screened for colour blindness using the Ishihara test [16]. The experimenter then introduced the experiment using a predefined script, and task instructions were given through a self-paced slide presentation. Following the instructions, participants completed a practice condition. All aspects of this practice condition were the same as the actual study conditions, except that the spatializations were replaced with images chosen from a set of 12 fruits (apple, banana, blueberries, cherries, kiwi, melon, nectarine, orange, pear, pineapple, starfruit, and strawberries). For the actual study, the six experimental conditions were then completed in random order.

5 RESULTS

Results were analyzed statistically using repeated measures analysis of variance (ANOVA) followed by Bonferroni-corrected pairwise comparisons. We first performed Q-Q plots to check our data distributions. Time data was log transformed so that it fit a normal curve. When Mauchly's Test of Sphericity indicated it was necessary, we used the Huynh-Feldt correction. Factors in the analysis were display type (3 levels), density (i.e. number of points, 2 levels), and seen (i.e. whether the images were new or seen in the training set, 2 levels).

Figures 3 and 4 show the response times and accuracy levels in different conditions, respectively. Column means for each of the conditions are shown on the bars within each figure. Error bars represent 95% confidence intervals.

5.1 Response Time

We expected most people to answer the yes / no questions quickly in all conditions, and were not surprised that most conditions were not significantly different in response time. However, we did observe a main effect of seen ($F(1,29)=20.3$, $\eta_p^2=0.41$, $p<0.001$). Participants responded significantly faster for images that had been seen in the training set compared to images that were new.

5.2 Accuracy

For accuracy, we observed significant main effects of display type ($F(2,58)=7.4$, $\eta_p^2=0.2$, $p<0.002$), point density ($F(1,29)=14.8$, $\eta_p^2=0.34$, $p<0.002$), and seen ($F(1,29)=14.1$, $\eta_p^2=0.33$, $p<0.002$). We also observed significant interactions for display \times density ($F(2,58)=4.3$, $\eta_p^2=0.13$, $p<0.019$) and display \times seen ($F(2,58)=3.2$, $\eta_p^2=0.1$, $p<0.049$).

Accuracy was significantly better overall when there were more points in the display, as shown by the higher green bars in Figure 4. Accuracy was also significantly better overall when the image was new. In other words, people were better able to recognize that they had not seen an image than that they had seen it. Post-hoc tests showed that both of these differences occurred for dots (density: $p<0.001$, seen: $p<0.027$) and 2-D landscapes (density: $p<0.008$, seen: $p<0.004$), but not for 3-D landscapes.

Overall, accuracy was 87.1% for dots, 85.3% for 3-D landscapes, and 79.1% for 2-D landscapes. Figure 5 illustrates these overall accuracy data. Both dots ($p<0.011$) and 3-D landscapes ($p<0.027$) had significantly higher accuracy than 2-D landscapes. However, this difference did not occur in all conditions, as shown in Table 1. Dots also had significantly higher accuracy than 3-D landscapes in the 1000-point condition.

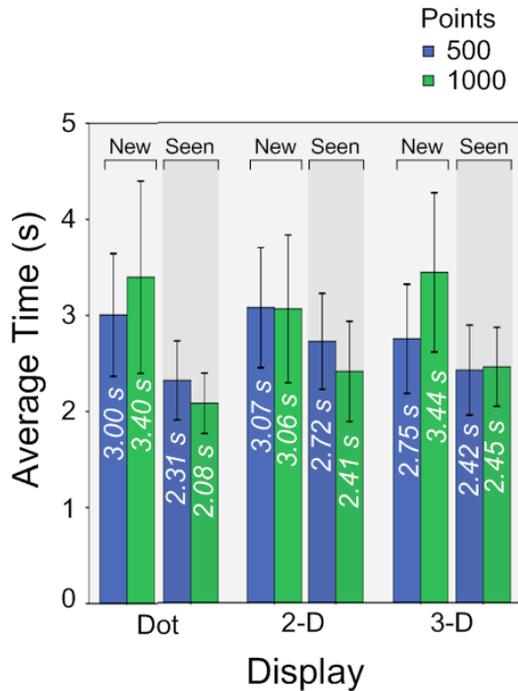


Figure 3: Response times for Dot, 2-D, and 3-D spatializations. Error bars show 95% confidence intervals.

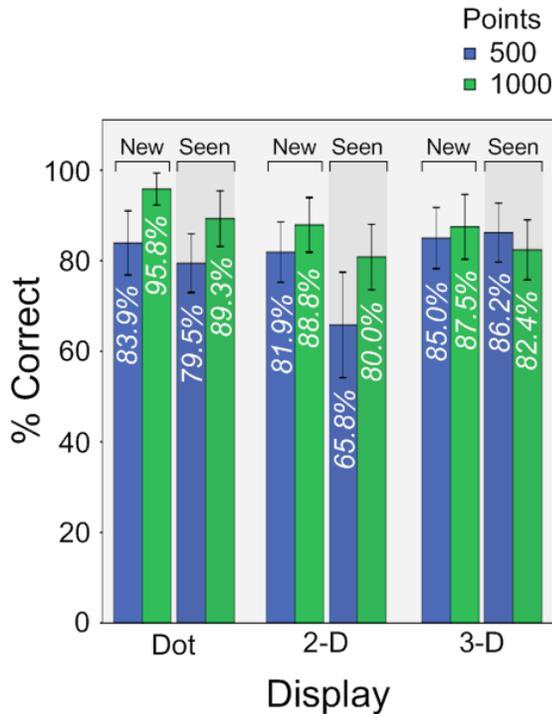


Figure 4: Accuracies for Dot, 2-D, and 3-D spatializations. Error bars show 95% confidence intervals.

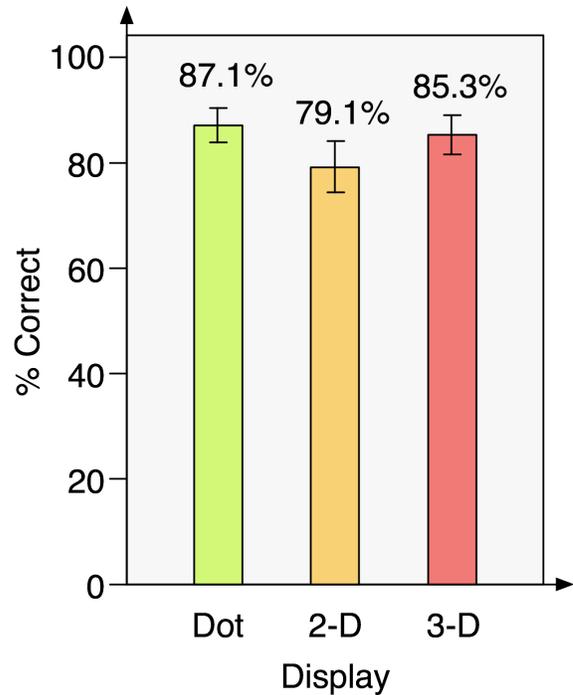


Figure 5: Overall accuracies for Dot, 2-D, and 3-D spatializations. Error bars show 95% confidence intervals.

Table 1: Interaction details for display × density and display × seen. Mean values and *p*-values are given for significant differences in memory accuracy.

Condition	Dot > 2-D	Dot > 3-D	3-D > 2-D
500 points			<i>Mn Δ%</i> = 11.7 <i>p</i> < .012
1000 points	<i>Mn Δ%</i> = 8.2 <i>p</i> < .004	<i>Mn Δ%</i> = 7.6 <i>p</i> < .011	
Seen image	<i>Mn Δ%</i> = 11.5 <i>p</i> < .021		<i>Mn Δ%</i> = 11.4 <i>p</i> < .017
New image			

6 DISCUSSION

We start our discussion by addressing the three key questions raised in Section 2:

Are landscapes or dots easier to remember? The most surprising result from this study was that memory was significantly more accurate with dot displays than with the 2-D or 3-D landscapes (refer to Figure 5). This contradicted our prediction that the abstraction provided by landscape displays would improve memorability. During pilot tests, the landscape displays even “seemed” easier to remember. Because response times were not statistically different between the Dot, 2-D, and 3-D factors, speed / accuracy trade-offs are an unlikely explanation for the higher accuracies observed for the Dot spatializations. One explanation for these results is that the contours and extra features within the 2-D and 3-D landscapes may actually be more visually distracting than helpful. Maybe people are able to sufficiently create abstract mental representations of Dot clusters without the visual abstractions provided by the landscapes. If such 2-D and 3-D visual landscape abstractions differ from a participant’s intuitive visceral mental model(s) based on Dot clusters, a slight confusion could occur and result in slightly lower accuracy scores. Another possibility is that the saturation-based coloring of the dot displays allowed lower density regions to fade into the background, producing a clear and memorable pattern of high density regions.

Does redundantly encoding data using colour and height improve memory compared to colour alone? Because height was the only difference between the 2-D and 3-D landscapes presented to participants, and the response times were not statistically different between the 2-D and 3-D landscapes, our results support this claim. The significantly higher accuracy results observed for the 3-D landscapes compared to the 2-D landscapes suggest that the redundant encoding of height improved memory.

It is interesting that accuracy was the worst overall with 2-D landscapes. If visual distraction was indeed the reason why dot displays were easiest to remember, one might expect that 3-D landscapes would be the hardest to remember since they present the most ‘distracting’ information. One possibility is that there was a greater range of brightness values in the 3-D images because of shading, which might make the images easier to remember. However, this is unlikely since the dot displays (which did not have 3-D shading information) had the best accuracy overall. Another possibility is that the discrete colour bands were distracting but the continuous 3-D shape was helpful in some way.

How is memory of spatialization influenced by the point density in the display? Figure 4 illustrates the significant and consistently higher accuracy results for the 1000-point density displays compared to the 500-point density displays. This result was also unexpected. Perhaps higher point densities may enable high-level features to emerge. With lower point densities, people may need to focus on a collection of smaller details within the displays instead of chunking a larger, more unified mental model. This conjecture is consistent with the different accuracy results obtained for the 3-D landscapes compared to the other conditions (see Figure 4). Specifically, accuracy results for the 500 vs. 1000 point densities were much more similar for the 3-D landscapes compared to the Dot and 2-D displays. Possibly, the height information within the 3-D landscape supports a more unified mental model of the data.

In addition to our three key questions, we were interested to observe that participants responded significantly faster to images that they had seen before in the training set compared to images that were new. Furthermore, for the Dot and 2-D displays, but not the 3-D display, participants were significantly more accurate correctly rejecting new displays compared to correctly accepting previously seen displays. The timing results are likely due to the extra uncertainty for the participants to consider whether a display (i) had been shown, but the participant forgot it, vs. (ii) had not been shown, and needed to identify it as new. The consistency of 3-D accuracy results may suggest that the 3-D displays were better able to provide sufficient features for participants to build a mental model of the display.

Considered together with previous work [10, 28], our results suggest that dot displays are equal to or superior to landscapes for most mental operations that have been studied, with the possible exception of density judgment [10]. This result is somewhat surprising – one might expect that the abstraction provided by landscapes would be useful, particularly for tasks involving memory. However, this does not appear to be the case, suggesting that designers should generally favour dot displays. If landscapes are to be used, preference for 2-D versus 3-D clearly depends on the task. While our study showed slight advantages for 3-D landscapes in terms of visual memory, studies with other tasks have shown better performance for 2-D landscapes [28] or mixed results depending on the task [10].

7 VISUALIZATION GUIDELINES

Our results are most relevant to designers of multidimensional visualization systems, particularly those involving information landscapes. Despite the popularity of 2-D and 3-D landscape displays, a growing body of empirical evidence suggests that dot-based displays may lead to better user performance, at least at many important low-level mental operations. Our visual memory results confirm the efficacy of dot-based spatializations and refute our prior

assumptions that 2-D and 3-D landscapes may be better for memory. Thus, we recommend that designers consider using dot-based spatializations unless there is some compelling reason to use a landscape display. One such reason might be user familiarity. For example, we are working with environmental simulation researchers who are accustomed to viewing contour plots and have been doing so for many years. In cases where a landscape display is chosen, we recommend that designers carefully consider the mental operations most important to the user’s task, and choose their design accordingly. In cases where users must frequently hold a landscape in memory for comparison with other landscapes, our results suggest a 3-D landscape may be better than a 2-D one.

8 FUTURE WORK

We did not expect to observe such pronounced differences between the 500 and 1000 point densities. Future work could focus on a user study containing a continuum of point densities. One could test for optimum densities of points for Dot-based, 2-D, and 3-D spatializations, and could also consider which type of display is best when the point density is so high that most points are visually obscured. Such studies could also compare coloring of points. For example, the study described in this paper had saturation gradient points as typically occur on Dot spatializations, and black points as typically occur on 2-D and 3-D spatializations. Future studies could explicitly compare the influences of these color choices. In addition, most research so far has used a bottom-up approach, focusing on which visualization techniques are best suited for various low-level mental operations. Further research should be done to consider the relative importance of these operations in more complex visualization tasks. For example, our study examined simple mental recall of landscape and dot images. Future work is needed to verify whether relative memorability of these displays changes during a higher-level data analysis task, where the image is not simply being remembered, but also interpreted.

9 CONCLUSIONS

We have demonstrated accuracy and timing differences for three main types of spatializations: dot displays, 2-D landscapes, and 3-D landscapes. Several useful results were observed that benefit visualization designers who are contemplating spatialization techniques to convey their data. Surprisingly, dot-displays afforded significantly higher accuracy results than 2-D or 3-D landscapes. We also observed that redundantly encoding data using colour and height in landscapes improved memory compared to colour alone, and that users were better able to remember denser spatializations. Finally, we quantified significant differences between a user’s ability to correctly recall a spatialization that they had seen before compared to a user’s ability to correctly identify a new spatialization. Collectively, these results help designers understand how to design individual spatializations; and, perhaps more importantly, how users are able to remember salient features while context switching between multiple different spatializations – a common need when analyzing multidimensional datasets.

ACKNOWLEDGMENTS

We thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding this research. We also thank Tamara Munzner for some helpful discussions leading up to this work.

REFERENCES

- [1] U. Bischoff, N. Diakopoulos, F. L6sch, and Y. Zhou: ThemeExplorer: A Tool for Understanding the History of the Field of Information Visualization. In InfoVisFun 2004, 2004.
- [2] U. Brandes and T. Willhalm: Visualization of bibliographic networks with a reshaped landscape metaphor. In Proc. Symp. Data Visualization (VisSym ’02), pp. 159-164, 2002.

- [3] C.M. Carswell and C.D. Wickens: Information integration and the object display: and interaction of task demands and display superiority. *Ergonomics*, 30(3), pp. 511-527, 1987.
- [4] M. Chalmers: Using a landscape metaphor to represent a corpus of documents. In Proc. European Conf. Spatial Information Theory (COSIT '93), Lecture Notes in Computer Science, 716, pp. 377-390, 1993.
- [5] A. Cockburn: Revisiting 2-D vs 3-D Implications on Spatial Memory. In Proc. 5th Australasian User Interface Conference (AUIIC 2004), pp. 25-31, 2004.
- [6] A. Cockburn and B. McKenzie: 3-D or not 3-D? Evaluating the Effect of the Third Dimension in a Document Management System. In Proc. ACM Conf. Human Factors in Computing Systems (CHI 2001), pp. 434-441, 2001.
- [7] A. Cockburn and B. McKenzie: Evaluating the Effectiveness of Spatial Memory in 2-D and 3-D Physical and Virtual Environments. In Proc. ACM Conf. Human Factors in Computing Systems (CHI 2002), pp. 203-210, 2002.
- [8] T. Cribbin and C. Chen: Visual-spatial exploration of thematic spaces: a comparative study of three visualisation models. *Electronic Imaging 2001: Visual Data Exploration and Analysis VIII*, pp. 199-209, 2001.
- [9] C.G. Christou, B.S. Tjan, and H.H. Bülthoff: Extrinsic cues aid shape recognition from novel viewpoints. *J. Vision*, 3, pp. 183-198, 2003.
- [10] S.I. Fabrikant: Spatial metaphors for browsing large data archives, Ph.D. thesis, University of Colorado-Boulder, Boulder, CO, 2000.
- [11] S.I. Fabrikant, D.R. Montello, and D.M. Mark: The distance-similarity metaphor in region-display spatializations. *IEEE Computer Graphics and Applications*, 26(4), pp. 34-44, 2006.
- [12] S.I. Fabrikant, D.R. Montello, M. Ruocco, and R.S. Middleton: The distance-similarity metaphor in network-display spatializations. *Cartography and Geographic Information Science*, 31(4), pp. 237-252, 2004.
- [13] M. Granitzer, W. Kienreich, V. Sabol, K. Andrews, and W. Klieber: Evaluating a system for interactive exploration of large, hierarchically structured document repositories. In Proc. IEEE Symp. Information Visualization (InfoVis'04), pp. 127-134, 2004.
- [14] E. Hetzler and A. Turner: Analysis experiences using information visualization. *IEEE Computer Graphics and Applications*, 24(5), pp. 22-26, 2004.
- [15] K. Hornbæk and E. Frøkjær: Do thematic maps improve information retrieval? In Proc. 7th IFIP Conf. Human-Computer Interaction (INTERACT'99), pp. 179-186, 1999.
- [16] S. Ishihara: Tests for Colour-blindness, Kanehara Shuppan, 1977.
- [17] H. Lam, R.A. Rensink, and T. Munzner: Effects of 2-D Geometric Transformations on Visual Memory. In Proc. 3rd Symp. Applied Perception in Graphics and Visualization (APVG '06), pp. 119-126, 2006.
- [18] K. Misue, P. Eades, W. Lai, K. Sugiyama: Layout Adjustment and the Mental Map. *J. Visual Languages and Computing*, 6, pp. 183-210, 1995.
- [19] D.R. Montello, S.I. Fabrikant, M. Ruocco, and R.S. Middleton: Testing the First Law of Cognitive Geography on Point-Display Spatializations. In Proc. Conf. Spatial Information Theory (COSIT '03), pp. 316-331, 2003.
- [20] G.B. Newby: Empirical study of a 3-D visualization for information retrieval tasks. *J. Intelligent Information Systems*, 18(1), pp. 31-53, 2002.
- [21] L. Nowell, E. Hezler, and T. Tanasse: Change Blindness in Information Visualization: A Case Study. In Proc. IEEE Symp. Information Visualization (InfoVis '01), pp. 15-22.
- [22] L. Nowell, R. Schulman, and D. Hix: Graphical Encoding for Information Visualization: An Empirical Study. In Proc. IEEE Symp. Information Visualization (InfoVis'02), pp. 43-50, 2002.
- [23] G. Robertson, M. Czerwinski, K. Larson, D. Robbins, D. Thiel, and M. Van Dantzich: Data Mountain: Using Spatial Memory for Document Management. In Proc. User Interface Software and Technology (UIST '98), pp. 153-162, 1998.
- [24] W. Schroeder, K. Martin, and W. Lorensen: The Visualization Toolkit. Prentice Hall PTR, second edition, 1998.
- [25] A. Skopik and C. Gutwin: Improving Revisitation in Fisheye Views with Visit Wear. In Proc. ACM Conf. Human Factors in Computing Systems (CHI 2005), pp. 771-780, 2005.
- [26] R. Spence: Information Visualization. Addison Wesley Publishers, 2000.
- [27] M. Tavanti and M. Lind: 2-D vs 3-D, Implications on Spatial Memory. In Proc. IEEE Symp. Information Visualization (InfoVis'01), pp. 139-145, 2001.
- [28] M. Tory, D.W. Sprague, F. Wu, W.Y. So, and T. Munzner: Spatialization Design: Comparing Points and Landscapes. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2007)*, 13(6), pp. 1262-1269, 2007.
- [29] S.J. Westerman and T. Cribbin: Mapping semantic information in virtual space: dimensions, variance and individual differences. *International J. Human Computer Studies*, 53(5), pp. 765-787, 2000.
- [30] M. Williams and T. Munzner: Steerable, progressive, multidimensional scaling. In Proc. IEEE Symp. Information Visualization (InfoVis'04), pp. 57-64, 2004.
- [31] J.A. Wise, J.J. Thomas, K. Pennock, D. Lantrip, M. Pottier, A. Schur, and V. Crow: Visualizing the non-visual: spatial analysis and interaction with information from text documents. In Proc. IEEE Symp. Information Visualization (InfoVis'95), pp. 51-58, 1995.