

A Comparison of Animated Maps with Static Small-Multiple Maps for Visually Identifying Space-Time Clusters

Amy L. Griffin,* Alan M. MacEachren,[†] Frank Hardisty,[‡] Erik Steiner,[§] and Bonan Li**

*School of Physical, Environmental and Mathematical Sciences, University of New South Wales-ADFA

[†]GeoVISTA Center, Department of Geography, The Pennsylvania State University

[‡]Department of Geography, University of South Carolina

[§]InfoGraphics Laboratory, Department of Geography, University of Oregon

**Department of Information Technology, Richland County Government

Although animated maps are widely promoted as ideal vehicles for learning and scientific discovery, there has been little empirical work that demonstrates their relative effectiveness in relation to static small-multiple alternatives. In this article, we attempt to clarify the issues related to the potential of animation from an explicitly geographic perspective, but one that is also grounded in broader cognitive science and human-computer interaction considerations. We compared the effectiveness of animated with static small-multiple maps, specifically in relation to map readers' ability to identify clusters that move over space and through time. In this study, we focused on several factors that might impact (or help explain) map readers' ability to correctly identify clusters. These factors included animation pace, cluster coherence, and gender. We found that map readers answer more quickly and identify more patterns correctly when using animated maps than when using static small-multiple maps. We also found that pace and cluster coherence interact so that different paces are more effective for identifying certain types of clusters (none vs. subtle vs. strong), and that there are some gender differences in the animated condition. This study is one of a small number of controlled experiments directed to the relative advantages of animated and static small-multiple maps. It provides the basis for further research that is needed to better understand the cognitive load involved in reading animated maps, to better describe and understand gender differences, and to investigate the efficacy of animated maps for other types of map reading tasks. *Key Words: map animation, small-multiple maps, visual cluster detection.*

Although the first animated maps were created as early as 1940 (Peterson 1995), they did not become a significant part of the spectrum of maps to which most people are exposed until the advent of the Internet as a medium for distributing maps in the 1990s (Harrower 2004). Animations in general and animated maps in particular are now commonly used as aids for teaching and learning about a wide variety of earth and social science processes in multimedia environments (e.g., Acevedo and Masuoka 1997; Blok et al. 1999; Lowe 1999; Harrower, MacEachren, and Griffin 2000; Slocum et al. 2000; Harrower 2002).

Animated maps have been proposed as ideal vehicles for learning and scientific discovery because they can explicitly represent dynamic systems and processes. Conventional wisdom holds that because animations explicitly depict temporal changes, they should more easily enable individuals to understand dynamic processes. Tversky, Morrison, and Betrancourt (2002) suggest that according to the Congruence Principle (i.e., the natural cognitive correspondence between an effective

graphic and mental representations of that graphic), it would not be surprising if the mental representations that people form of dynamic processes shared structural characteristics with animations (i.e., that people mentally represent dynamic processes as animations). However, there is evidence to suggest that people mentally represent dynamic processes or systems through a series of static small-multiple snapshots based on critical moments (i.e., moments of important change) rather than as a dynamic representation (Hegarty, Kriz, and Cate 2003; Lee, Klippel, and Tappe 2003).

Despite wide promotion of the benefits of animation, there is little empirical evidence demonstrating that animated representations are indeed superior to static small-multiple representations of processes. Studies of static small-multiple and animated cartographic representations have produced mixed results. Kossoulakou and Kraak (1992) compared static small-multiple and animated forms of several different types of cartographic symbolization. They found that participants were able to answer questions more quickly when they used the

animated representation, but found no differences in the number of correct answers by representation type. In contrast, Cutler (1998) found that children answered significantly fewer questions correctly and took longer to answer in an experiment that compared animated isarithmic maps with static small-multiple maps. In a recent qualitative evaluation of software that provided map readers with both animated maps and static small-multiple maps, Slocum et al. (2004) found that participants felt that animations and static small-multiples were best used for completing different tasks: animations for identifying general trends and static small-multiples for comparing specific time points.

In a recent review of noninteractive animations¹ that did not include any cartographic representations, Tversky, Morrison, and Betrancourt (2002) found experimental design flaws in more than twenty studies that purport to show a benefit of animated representations over static small-multiple representations, and concluded that there was no evidence to suggest that animated representations help people learn more effectively. They found two main types of design flaws in comparative experiments: (1) the animated and static small-multiple graphics contained incomparable levels of information (i.e., that because the animated graphics contained microsteps, they showed more information than the static small-multiple graphics); and (2) many studies used different procedures while testing the animated and static small-multiple graphics (e.g., the animated graphics were interactive, the static small-multiple graphics were not). The presence of these experimental flaws led Tversky, Morrison, and Betrancourt to conclude that the differences the studies found may have been due to the procedures used rather than to the graphics themselves.

One explanation for the better performance of static small-multiple representations compared with animated representations may lie in the perceptual and cognitive processes involved in processing the information contained in visual representations. A number of authors have noted that the extra cognitive effort required to interpret animated representations (compared with static small-multiple representations) may impede readers' ability to efficiently extract information from them (Tversky, Morrison, and Betrancourt 2002; Hegarty, Kriz, and Cate 2003; Hegarty 2004; Price 2004). Two commonly cited potential factors relate to the (sometimes) ephemeral nature of features in animations that may lead to an overload of working memory and to how the viewer directs his or her attention. In many cases both of these problems could be attenuated by incorporating interactivity into the animated representations.

However, we believe it is important to achieve some understanding of factors that may affect the perceptual and cognitive processes that are involved in extracting information from *noninteractive* animated representations, as this knowledge could help guide and improve the design of both interactive and noninteractive animations.

Tversky, Morrison, and Betrancourt's (2002) review focused on animated representations that were designed to help students understand nongeographic dynamic processes and systems, and did not include animated representations designed for other purposes (the examples they reviewed included diagrams that explain the working of the circulatory system, how a bicycle pump works, and Newton's laws of motion).

Another purpose for which animated representations might be used is for knowledge construction and discovery while viewers are exploring data. In this case, the viewer's primary goal might be to notice a pattern in his or her data, rather than (at least initially) trying to understand the process that is creating that pattern. Our goal in this experiment was to investigate the relative effectiveness of animated and static small-multiple representations of space-time information in aiding viewers with knowledge construction and discovery tasks.

One case in which we believed that animated maps would be a more effective tool than static small-multiple maps was in detecting clusters that move through space and over time.² Visual cluster detection is essentially a process of identifying one or more areas that are internally homogenous but different from their surroundings. Indeed, Gestalt psychologists identified similarity as one of the factors that aids the perceptual grouping process (MacEachren 1995).

The study presented here is a test of a specific hypothesis, that an animated representation will be more effective than a static small-multiple representation for visually detecting moving clusters. The rationale for this hypothesis is the assumption that animation takes advantage of an additional Gestalt principle: common fate. This principle holds that objects that are moving together (i.e., in the same direction at the same rate) are seen as a group. In particular, we thought that the apparent motion inherent in an animated presentation (of the individual scenes making up the static small-multiple representation alternative) would help in identifying clusters that may be of interest but that are perhaps more difficult to identify from multiple, juxtaposed frames.

In addition to comparing viewers' abilities to identify moving clusters from a noninteractive animated map with the ability to identify the same moving clusters from a set of static small-multiple maps, in this study we also

investigated how two factors (animation pace and cluster coherence) might influence users' ability to identify moving clusters.

In the section that follows, we outline the methodology for our study. Then we provide detailed results of the experiment. The final section offers a discussion of the results, our conclusions, and suggestions for future research.

Methodology

The study attempts to address some of the concerns raised by Tversky, Morrison, and Betrancourt (2002) about previous experiments by developing a carefully controlled experiment in which the animated and static small-multiple map representations to be compared differ only in respect to whether they are animated or not; thus other typical differences that exist between "real-world" animated and static small-multiple maps of time series data are removed. We also controlled for other factors that might have an impact on the relative effectiveness of animated and static small-multiple map representations, but that are not part of the difference between these forms of representation (the kind of map, the structure of the geographic region represented, the nature of the data signified, etc.). To do so, we opted for maplike displays with constructed data sets in which we could control several aspects of the spatial and temporal characteristics of the data. Subsections below provide details on (a) the construction of the visual (maplike) stimuli, (b) the creation of the test instruments, (c) the experimental design, (d) the test participants, and (e) the experimental procedures.

Development of Moving Cluster Visual Stimuli

We developed a random pattern generation method to create a set of test stimuli that could be used to explore the relative effectiveness of animated and static small-multiple representations for allowing subjects to detect moving and growing patterns. The development of these pattern generation methods was itself a challenge. There is little or no previous work in the area; there are no generally accepted methods (such as statistical tests) for identifying such space-time patterns, and thus no standard way to measure space-time clustering within the patterns we generated. A method that addresses part of the problem is the scan statistic developed by Kulldorff (Kulldorff and Nagarwalla 1995), a method derived from early work by Openshaw on the Geographical Analysis Machine (GAM; see Openshaw, Charlton, and Craft 1988). However, the scan statistic

detects only space-time cylinders (if time is treated as a third dimension). Moving clusters can be thought of as space-time prisms; growing clusters as space-time cones. Neither is detected by the scan statistic.

Given the lack of a generation method for creating moving clusters or a statistic for assessing space-time clusters, we developed a cluster generation method that produces, on average, statistically significant spatial clusters at each step, where the location of these clusters changes systematically over time (we used a linear change in location over time in all cases). These simplifying assumptions make it possible to create a series of test stimuli for which we are confident that the space-time pattern is a kind of moving cluster. The moving clusters that this method generates are only a subset of the possible kinds of moving clusters that may exist in real-world data. Our clusters are all relatively compact, are all made up of cells that are identical in size and shape, and are each big enough to be individually identifiable at the display size and resolution we used ($1,280 \times 960$ on 17-inch monitors).

The maplike displays we generated are analogous to choropleth maps of data by census tract or county, without the variability of real places. A display consists of a 36×21 array of hexagons. Each hexagon could take on one of five different data categories represented by color values (see Figure 1). We used a heterogeneous Poisson process to generate a spatially clustered pattern at each time step. In other words, we created a random background pattern, and then within that an area of higher (but still randomly generated) intensity. We then forced the center of the pattern to move across the study area in a linear manner.

The intensity of the cluster was varied by increasing the color value of a number of the cells within the cluster area by one setting, allowing for cells to be repeatedly intensified. We defined moving clusters as circular areas with radii of four hexagonal shapes. We varied the proportion of intensified cells to produce subtle or strong moving clusters (see Figures 2 and 3). Because one of our

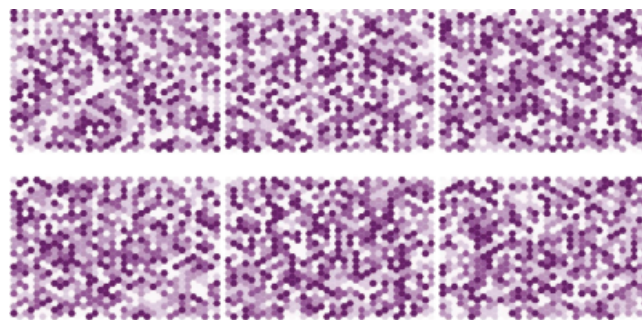


Figure 1. Frames from a trial with no moving cluster.

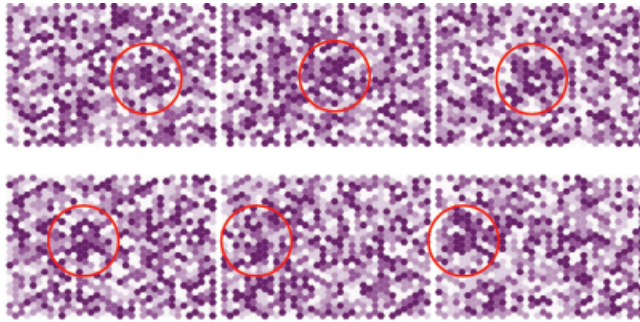


Figure 2. Frames from a trial with a subtle moving cluster (moving from the right-center toward the left).

goals in creating these patterns was to produce a range of possible time-space patterns for which parameters of the representation might have an impact, we did pilot studies that helped to calibrate the match between statistical measures and perception. In other words, our pilot studies identified patterns that were easily, but not always, discerned (our strong clusters) as well as patterns that were typically discernable, but difficult to discern (our subtle clusters).

The Java code we developed for generating the test patterns is a small but measurable addition to the set of methods available for manipulating hexagonal data structures, following Dan Carr's work (Carr, Olsen, and White 1992). The code is available in both source and executable forms from <http://people.cas.sc.edu/hardistf/hexamap>. The features of the code include a number of routines for manipulating hexagonal arrays and output in both XML and SaTScan forms. We have released the code under an Open Source license (LGPL) that permits both modification and redistribution. The web site includes executable versions for Microsoft Windows, Macintosh OS X, Sun Solaris, and GNU Linux. By providing this code and the executables, we make it practical for others to critique or extend the research presented here by exploring in detail the cluster generation method and by modifying that method to generate

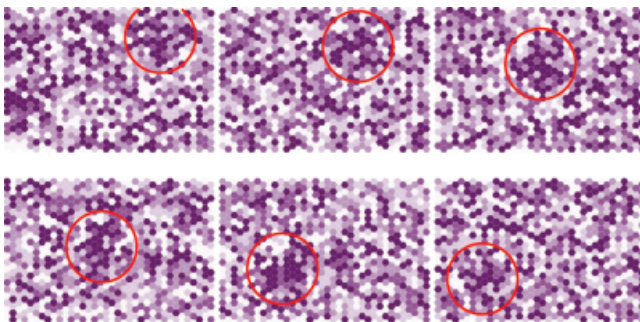


Figure 3. Frames from a trial with a strong moving cluster (moving from the upper-right toward the lower-left).

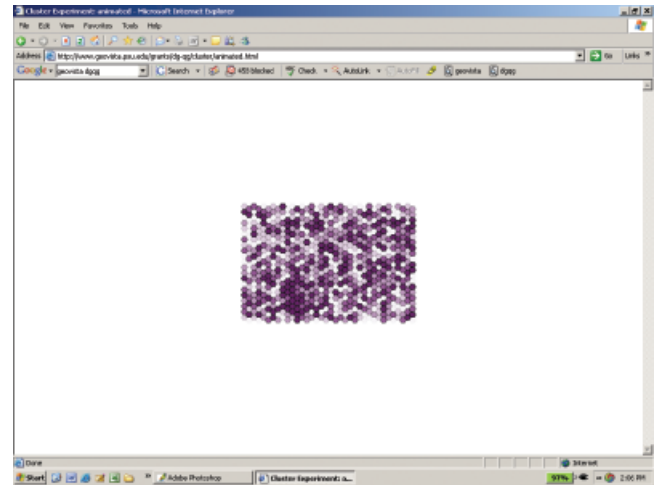


Figure 4. Test instrument appearance: Example of a frame from the visual stimulus used in the animated condition. The animated version of the test instrument can be viewed at: <http://www.geovista.psu.edu/grants/dg-qg/cluster/animated.html>.

alternative kinds of moving clusters for subsequent testing.

Test Instruments

We created two versions of the test instrument, each containing identical visual information. Both versions were created using Macromedia Flash software. One version contained a map animation with six frames (the "animated condition"; see Figure 4); the other displayed a static small-multiple array of the same six frames (the "small-multiple condition"; see Figure 5).

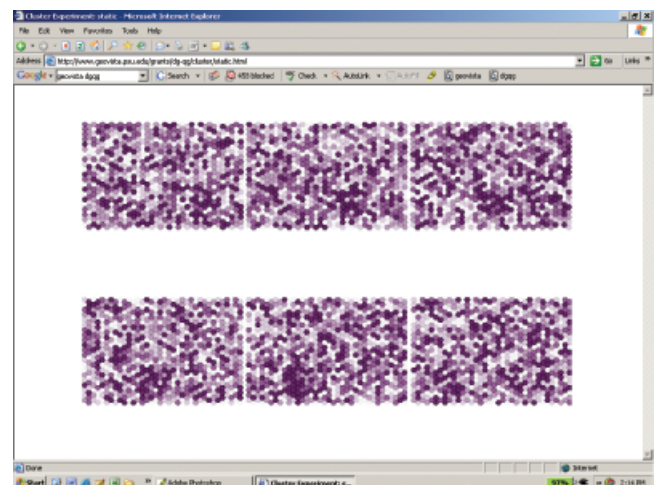


Figure 5. Test instrument appearance: Example of a visual stimulus used in the small-multiple condition. The static small-multiple version of the test instrument can be viewed at: <http://www.geovista.psu.edu/grants/dg-qg/cluster/static.html>.

In the animated condition, participants saw each six-frame animation three times. Each run of the animation was separated from the next by a gray screen that flashed to indicate that the animation was starting again. To control for the possibility that there might be a positive relationship between a map reader's ability to identify moving clusters and the size of the visual stimuli (the map size), we used visual stimuli of the same size in both the animated and small-multiple conditions.

As described above, we constructed the visual patterns from regular arrays of hexagons. We chose to use an abstract shape for the visual stimuli (as opposed to real geographic areas) in order to remove familiarity with particular geographic places as a factor in the experiment and to eliminate the possibility that the shape and size of geographic units could have an impact on inferences about the pattern (or lack of pattern) that participants were seeing. In the future this work could be extended to real places (e.g., counties or local government areas) using the same experimental framework to determine what effect (if any) place-based knowledge or varying the size and shape of geographic units has on the patterns people see or do not see.

Experimental Design

To compare animated maps with the static small-multiple maps, we used a within-subjects experimental design in which each participant saw both the animated and the small-multiple condition.

In this experiment, our goal was to compare animated with static small-multiple representations across a range of data and display conditions. To make the experiment practical, we had to control for many factors in order to test a few. Pilot studies and observations about use of the many different map animations we have implemented in the past led us to focus on two factors that we expected to have an impact on the relative advantages of each type of representation. These factors are (a) the coherence of space-time patterns (pilot studies suggested that a difference between the two map forms is most pronounced for relatively difficult judgments—in situations in which the presence of a moving cluster is not obvious); and (b) the pace of the animation (work with several map-based animations convinced us that there is probably an “optimal” pace for at least some map animations that can result in visual “popout” effects that lead to nonobvious space-time patterns becoming obvious). As a result, we designed the experiment to test for these two factors while controlling for others. The experiment tested each participant on three distinct coherence levels (no pattern, subtle pattern, and

strong pattern) and at four distinct paces (5, 7, 9, and 11 sec).

This led to an experiment with twelve conditions, each of which was included two times within the test instrument (i.e., there were twenty-four trials, two for each pattern coherence/pace combination). Due to our desire to test for both pattern coherence and animation pace while keeping the number of study participants to a reasonable level, we used a balanced Latin square design within the test instrument for each condition to control for order effects related to pace and coherence. In a balanced Latin square design, each condition precedes and follows each other treatment equally often (e.g., the combination of pace = 5 sec, coherence = no pattern precedes the combination of pace = 5 sec, coherence = subtle pattern as often as it follows it), and each condition appears in each ordinal position the same number of times. In other words, within the total population of experiment participants, a particular visual stimulus combination (e.g., pace = 5 sec, coherence = no pattern) will appear once in each trial position.

Participants

Twenty-four students (twelve males and twelve females), who were mostly between the ages of eighteen and twenty-four years of age (twenty-two of the twenty-four participants), and who were from varied academic backgrounds, participated in the experiment. They were paid a small sum of money in return for their participation. As we did not want to bias our sample by including a high proportion of participants who had extensive experience with animated maps, we specifically avoided recruiting students who had taken the department's introductory cartography or dynamic cartography courses.

Experimental Procedures

We carried out the experiment in a department computer lab. We counterbalanced the animated and small-multiple conditions so that half of the participants saw the animated condition first, and the other participants saw the small-multiple condition first. We also controlled for gender in this counterbalancing, but did not counterbalance participants by age because almost all participants were in the same age class.

Participants in both conditions initially saw a welcome page that explained the purpose of the experiment and described the task that they would be required to perform: identifying whether or not a moving cluster was

present in the display. In the small-multiple condition, this welcome screen also included a diagram that showed the order in which the screens should be viewed (i.e., the top row left to right, followed by the bottom row left to right). Next, participants in both conditions saw an example of a trial that contained a moving cluster. After seeing the visual stimulus, the participants were asked to record whether they saw a cluster or not by clicking either a button labeled “moving cluster” or another labeled “no cluster.” In addition to the participants’ answers, the computer also recorded the amount of time it took for participants to respond by clicking on one of the two buttons. After they indicated whether or not they had seen a moving cluster, the participants were told whether they had correctly identified the pattern, and were shown an example of a trial that did not contain a moving cluster.

After seeing the examples, participants completed twenty-six trials. The first two trials used the same strong (i.e., high coherence) cluster patterns for each participant and were not analyzed, leaving twenty-four trials of actual data for each participant. These practice trials helped participants to orient themselves to the task. Across the twenty-four core trials, we presented participants with stimuli having different levels of pattern coherence for different lengths of time (i.e., paces), using the Latin Squares procedure detailed above to balance the order of presentation. In the animated condition, the four different animation paces meant that for individual trials, participants had different total lengths of time to view and interpret the particular display sequence of six maps contained in that trial. To remove this as a factor in comparing the animated and small-multiple conditions, we adjusted the display times for the static small-multiple stimuli in the same way (i.e., we created a “pace” for the static small-multiple maps). For each individual trial, pace in the animated condition was matched to viewing time for the small-multiple condition.

At the end of the session, participants filled out a demographic questionnaire where they indicated their age, gender, field of study, and certainty in their answers on a Likert scale that ranged from 0 to 100. There was also an opportunity for participants to make comments about the experiment. After completing the procedure for the initial condition to which they were assigned (animated or small-multiple), participants worked with the test instrument for the other condition.

Results

The test instruments collected two pieces of data for each trial that each participant completed: the answer

he or she recorded and the time he or she took to answer. We analyzed the data separately for each condition, then compared the animated with the small-multiple conditions. In analyzing the animated and small-multiple conditions separately, we used the chi-square statistic to look at the number of correct answers under different experimental conditions (i.e., pace and cluster coherence) as well as different participant demographic groupings, and either ANOVA (analysis of variance) or a two-sample *t*-test to examine the times under different experimental conditions. To compare the conditions, we used the McNemar test to examine the number of correct answers in each condition, and the Wilcoxon sign-rank test to look at time differences. We first report on the comparison of the animated and small-multiple conditions and then report the results for each condition separately. To understand how variables related to each other, we also stratified trials by variables and looked at the effects of other variables on each of the strata.

Comparison of the Animated and Small-Multiple Conditions

Participants identified patterns correctly more often using the animation than they did using the small-multiples (70.5 vs. 64.8 percent; $\chi^2 = 4.80$, $p < 0.03$). They also took significantly less time to answer using the animation (1.1 vs. 1.4 sec; $z = -7.90$, $p < 0.01$).

When we looked at the number of patterns that participants identified correctly at different paces in the two conditions, we found that there was no difference between the animation and the small-multiple representation at either the slowest or fastest pace (see Figure 6). However, patterns were identified correctly significantly more often in the animation at moderate paces (7 sec: 72.2 vs. 61.1 percent, $\chi^2 = 3.80$, $p < 0.05$; 9 sec: 79.2 vs. 68.8 percent, $\chi^2 = 4.40$, $p < 0.04$). Participants also identified patterns significantly more quickly using the animation, with the difference in the time it took to answer increasing substantially at slower paces (see Figure 7). Interestingly, in both conditions, participants took longer (although not significantly longer) to answer at the slowest pace, perhaps suggesting that when the pace is too slow the visual system cannot maintain the Gestalt grouping of common fate.

Although there were no differences in the number of patterns identified correctly between the animated and small-multiple conditions at different levels of pattern coherence, there were significant differences in the amount of time it took to identify patterns (no cluster: $z = -3.70$, $p < 0.01$; subtle cluster: $z = -5.00$, $p < 0.01$; strong cluster: $z = -5.00$, $p < 0.01$). Animation provided

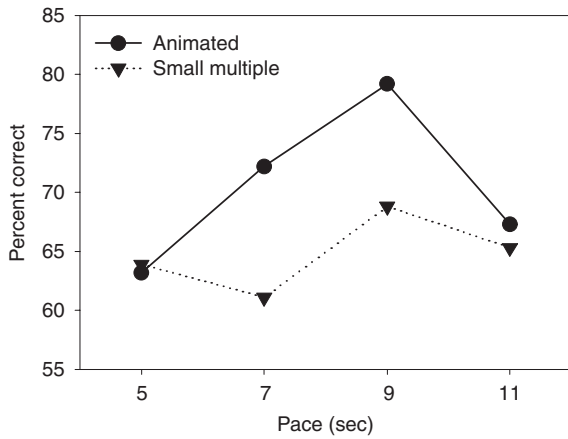


Figure 6. Percentage of patterns correctly identified in the animated and small-multiple conditions, stratified by pace.

the strongest advantage in cases where the trials contained subtle clusters (see Figure 8).

There were significant differences between the animated and small-multiple conditions in the number of correct responses provided by males when responses were stratified by gender. Males identified patterns correctly more often in the animated condition than in the small-multiple condition (74.3 vs. 66.3 percent; $\chi^2 = 4.80$, $p < 0.03$). There was no significant difference for females (66.7 vs. 63.2 percent; $\chi^2 = 0.72$, $p < 0.40$). Both males and females were significantly faster in identifying patterns with the animation than with the small-multiples (males: $z = -6.70$, $p < 0.01$; females: $z = -4.50$, $p < 0.01$). The difference between the two conditions was larger for males than for females (0.3 sec vs. 0.2 sec).

Animated Condition

Overall, participants correctly identified patterns (or lack of pattern) in 70.5 percent of trials ($n = 576$) in the

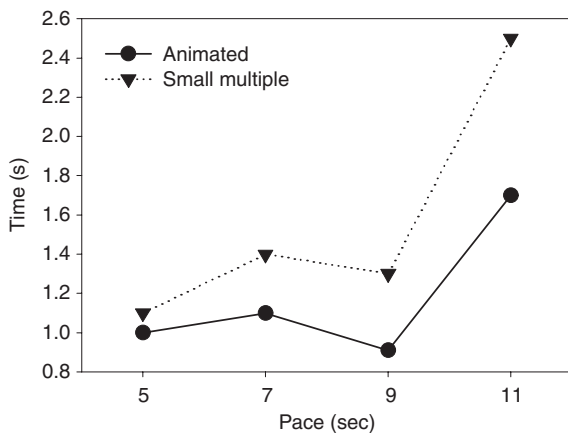


Figure 7. Time taken to answer in both the animated and small-multiple conditions, stratified by pace.

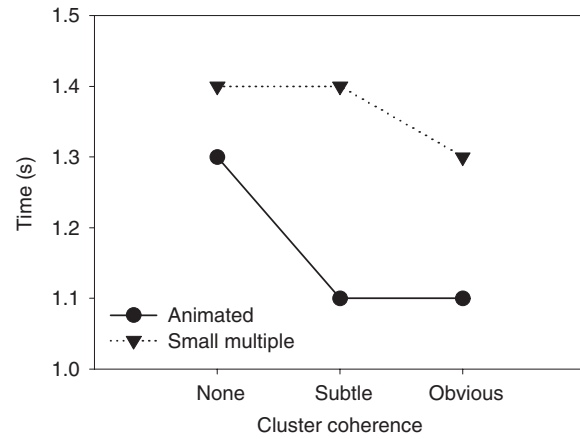


Figure 8. Time taken to identify patterns in both the animated and small-multiple conditions, stratified by coherence.

animated condition. It took participants significantly less time when they identified a pattern correctly than when they did not identify a pattern correctly (1.1 vs. 1.3 sec; $z = -3.93$, $p < 0.01$).

Pace and the Animated Condition. As the pace of the animation decreased (from 5 sec to 11 sec), participants identified a higher number of patterns correctly ($\chi^2 = 24.1$, $p < 0.02$), except at the slowest pace, where there were fewer correct answers than in all but the fastest pace:

- 5 sec (fast) 63.2 percent correct
- 7 sec 72.2 percent correct
- 9 sec 79.2 percent correct
- 11 sec (slow) 67.4 percent correct.

These results imply that too little viewing time per frame makes it hard to process the scene; as pace slows down, performance improves up to the point at which the visual system cannot maintain the Gestalt grouping principle of common fate. There were no significant differences in the time it took participants to answer at different animation paces.

When we stratified the trials by cluster coherence, we found that there was an interaction of pace with cluster coherence. For trials with no cluster or a strong cluster, the second fastest pace (7 sec) was the pace at which participants identified the most patterns correctly. However, trials with subtle clusters were best identified at a slower pace (9 sec; see Figure 9). These differences between cluster coherence types were significant at 7 sec ($\chi^2 = 18.2$, $p < 0.01$) and 11 sec ($\chi^2 = 13.4$, $p < 0.01$). The subtle clusters require a slower pace—thus more time per view to process—up to the point when the

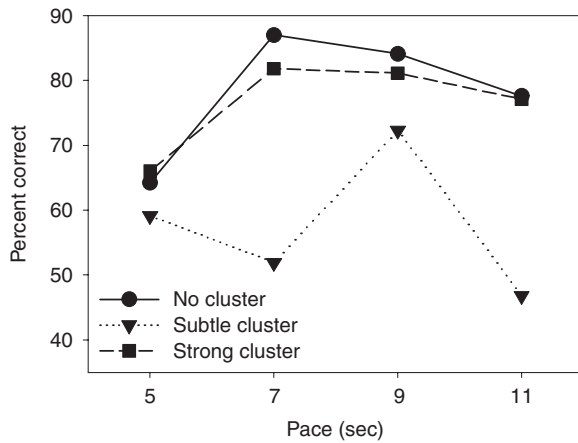


Figure 9. Percentage of patterns correctly identified in the animated condition at different paces, stratified by cluster coherence.

increased time between frames makes it hard to maintain short-term memory of what was present in the previous frame (i.e., the appearance of common fate). Additionally, there appears to be a range of effective paces across which animations can be successfully used to identify patterns, but this range is perhaps narrower for subtle patterns than for strong patterns. These results suggest that providing interactive control of animations is likely to be critical to the effective use of pace as a dynamic variable, as it will be difficult or impossible to predetermine what pace will work for which dataset.

Cluster Coherence and the Animated Condition. Not surprisingly, we found that there were significant differences ($\chi^2 = 24.1, p < 0.01$) in the number of patterns that participants correctly identified in trials with different levels of cluster coherence. Participants were better at identifying strong moving clusters and patterns with no moving cluster than patterns with subtle moving clusters:

- No pattern 77.6 percent correct
- Subtle pattern 57.3 percent correct
- Strong pattern 76.6 percent correct.

It took participants significantly longer to answer in trials where there was no cluster than in trials where there was a strong cluster (1.3 vs. 1.1 sec; $F = 4.20, p < 0.02$, Bonferroni corrected). This result is one that we would expect; that is, the no-cluster trials require an exhaustive visual search, while trials that contain clusters are self-terminating visual searches. There were no significant differences in the time it took to answer in other group comparisons.

Gender Effects and the Animated Condition. Males were significantly more likely to identify patterns correctly than were females ($\chi^2 = 4.04, p < 0.04$):

- Males 74.3 percent correct
- Females 66.7 percent correct.

This result is unusual, for although modest gender differences have been commonly reported for some map-reading tasks associated with navigation (e.g., Montello et al. 1999; Malinowski and Gillespie 2001; Lloyd and Bunch 2005), they have not been commonly reported for thematic map reading (e.g., Gilmartin and Patton 1984; Evans 1997). There was no significant difference in the amount of time it took males or females to identify patterns.

We also stratified the trials by cluster coherence and animation pace to investigate whether there were other experimental factors that are important for describing this difference between males and females. Pattern coherence seems to be one factor that differentially affects the performance of males and females. There is a substantial, but not statistically significant difference between males' and females' abilities to correctly identify patterns in trials where there is either a subtle pattern or a strong pattern (subtle pattern: $\chi^2 = 2.13, p < 0.15$; strong pattern: $\chi^2 = 3.51, p < 0.06$; see Figure 10). This pattern of males identifying more patterns correctly is much less marked in the trials that did not contain moving clusters ($\chi^2 = 0.03, p < 0.86$).

The differential rate at which males and females correctly identified patterns also exists at faster animation paces. As pace decreases, males have increasing numbers of correct pattern identifications until the slowest pace, which is the most inaccurate for them (see Figure 11). For males, these performance differences

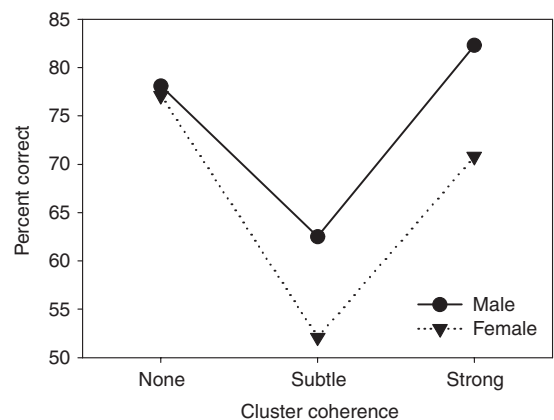


Figure 10. Percentage of patterns identified correctly in the animated condition at different levels of pattern coherence, stratified by gender.

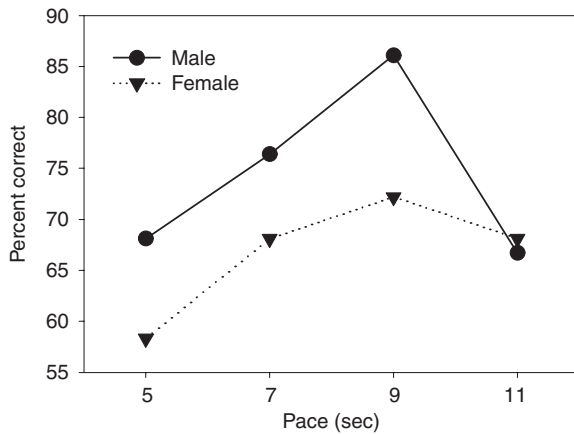


Figure 11. Percentage of patterns identified correctly in the animated condition at different paces, stratified by gender.

(between paces) are significant ($\chi^2 = 9.10$, $p < 0.03$). Females have a similar overall pattern, except that they have fewer correct responses than males at all but the slowest pace, and the magnitude of the pace effect is weaker and not statistically significant ($\chi^2 = 3.40$, $p < 0.34$). Males were significantly better than females at identifying patterns for pace three ($\chi^2 = 4.20$, $p < 0.04$); there was no statistically significant difference between males and females at any other pace.

Small-Multiple Condition

Overall, participants correctly identified patterns from the small-multiple representation trials 64.8 percent of the time ($n = 576$). As in the animated condition, participants took less time to correctly identify a pattern than to incorrectly identify patterns (1.3 vs. 1.4 sec; $z = -2.80$, $p < 0.01$).

Pace and the Small-Multiple Condition. In the small-multiple condition, the term “pace” refers to the length of time that the participants saw the representation. The paces used in this condition were identical in length to those used in the animated condition. Although there was no significant difference in the number of correct responses between the paces in the small-multiple condition (as we observed in the animated condition; $\chi^2 = 2.00$, $p < 0.57$), there does seem to be a threshold between the two fastest paces and the two slowest paces (i.e., between 7 and 9 sec):

- 5 sec (fast) 62.7 percent correct
- 7 sec 61.6 percent correct
- 9 sec 68.8 percent correct
- 11 sec (slow) 66.0 percent correct.

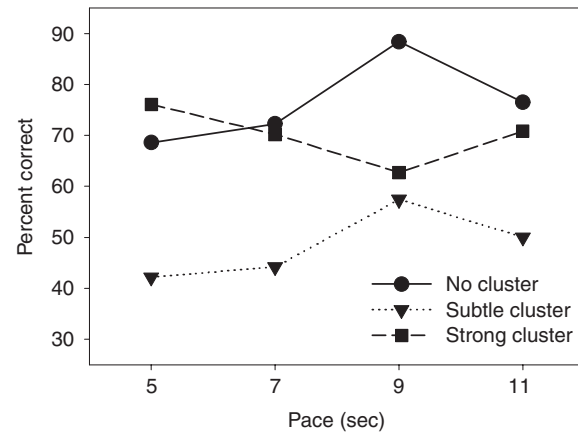


Figure 12. Percentage of patterns correctly identified in the small-multiple condition at different paces, stratified by cluster coherence.

An explanation for this pattern is simply that a display must be visible long enough for the participant to scan all of the small-multiples. This hypothesis could be tested by comparing the ability of participants to correctly identify patterns from displays with increasing numbers of small-multiples. We would expect that as the number of multiples increases, so too would the amount of time needed to correctly identify a particular number of patterns.

There was a strong interaction between cluster coherence and the pace at which the small-multiple representations were displayed. There was a statistically significant difference between the levels of pattern coherence for all paces (5 sec: $\chi^2 = 12.4$, $p < 0.01$; 7 sec: $\chi^2 = 10.4$, $p < 0.01$; 9 sec: $\chi^2 = 11.4$, $p < 0.01$; 11 sec: $\chi^2 = 8.40$, $p < 0.02$). For trials that contained either no cluster or a subtle cluster, participants identified the most patterns correctly at the second slowest pace (9 sec). In contrast, for trials that contained a strong cluster, the fastest pace (5 sec) was most conducive to participants correctly identifying patterns (see Figure 12). One potential explanation for these results is that the faster pace that produced better identification of strong clusters forced participants to rely (unconsciously) on preattentive visual processing instead of focused attention for determining if a pattern was present or not. In contrast, preattentive vision may not have been sufficient to recognize subtle clusters or lack of clusters.

Cluster Coherence and the Small-Multiple Condition. As in the animated condition, there was a significant difference in the number of correct responses for different coherence levels ($\chi^2 = 35.4$, $p < 0.01$).

Participants were better at identifying patterns when they contained either no cluster or a strong cluster:

- No cluster 76.0 percent correct
- Subtle cluster 48.4 percent correct
- Strong cluster 69.8 percent correct.

Gender Effects and the Small-Multiple Condition. In the small-multiple condition, there was no statistically significant difference in the number of correct answers between males and females ($\chi^2 = 0.61, p < 0.43$). However, females were slightly faster at identifying patterns than males (1.3 vs. 1.4 sec; $z = -2.80, p < 0.01$). There were also no significant interactions of gender with pace or cluster coherence.

Discussion and Conclusions

Overall, our results show that the animated representation enabled users to more often correctly identify whether a particular type of pattern was present than did the static small-multiple representation. Participants were also able to come to a conclusion about what they saw more quickly while viewing an animated representation.

Our results do not provide any evidence to support the often-made contention that animations require extra cognitive effort for information extraction when compared with static small-multiple representations. It may be that the cognitive load associated with visually identifying moving clusters is relatively light, and those factors that influence the map reader's perceptual capabilities are more important for successfully completing this task.

Performing a cognitive demands analysis, a technique described by T. Green (1989), for the task we studied (i.e., visual identification of moving clusters) provides some support for our contention that animations should be (and are) more effective for this task than are static small-multiples. The visual identification of moving clusters is a special case of the more general problem of visually identifying clusters. In order to identify a cluster, map readers must first *notice* a graphical difference between different locations on a map. A prerequisite for noticing a difference in a visual representation then would be whether or not the difference is visible. Once map readers have noticed a difference, they then have to *identify* whether or not the difference is a cluster or some other type of graphical difference. In an experiment that asked map readers to outline clusters from graduated circle maps, Slocum (1983) found that two characteris-

tics of the maps (proximity and figure-ground relationships) influence whether map readers define particular groupings of map symbols as clusters. Finally, map readers may also *interpret* the identified cluster (e.g., decide whether or not it is likely to be a significant or a spurious cluster).

In the case of visually identifying moving clusters from static small-multiples, we would expect the map reader to iterate through the same three cognitive tasks we described above for each multiple contained in the collection of small-multiples (see Figure 13). In animations, by contrast, we suggest that this process of examining each multiple (or in this case, each frame) can be bypassed if the map reader focuses on the characteristics of the motion contained in the animation. In animated exemplars containing a moving cluster, there should be coherent motion rather than the random patterns of motion contained in exemplars that do not contain moving clusters. In other words, because animations enable map readers to focus on an *additional* graphical characteristic of the patterns (motion in addition to proximity and figure-ground relationships), map readers should be able to identify clusters both more easily and more quickly in the animated condition than they can in the small-multiple condition.

We did not collect any direct evidence of where participants were directing their attention (such as eye-movement measurements), and thus cannot conclusively comment on any differences that may have existed in where participants directed their attention in the different conditions (i.e., whether map readers viewed each of the small-multiple frames sequentially). However, it is plausible to speculate that map readers, through their prior experience of looking at animations (of any type), have (consciously or unconsciously) developed appropriate and effective strategies for directing their attention while viewing an animation (i.e., that they focus on the characteristics of the motion).

Our results do provide some evidence in support of this hypothesis in that the optimal pace for correctly identifying different types of patterns in the two conditions differed. In trials that contained a strong cluster in the small-multiple condition, participants correctly identified the highest number of patterns correctly at the fastest pace. It seems likely that for the small-multiple condition, participants directed their attention to each frame more or less sequentially while looking for a cluster. We would expect participants who were using this strategy to identify a higher number of patterns correctly at slower paces, as in the slower paces they would actually have enough time to fixate on and visually analyze each of the multiples. Indeed, we did

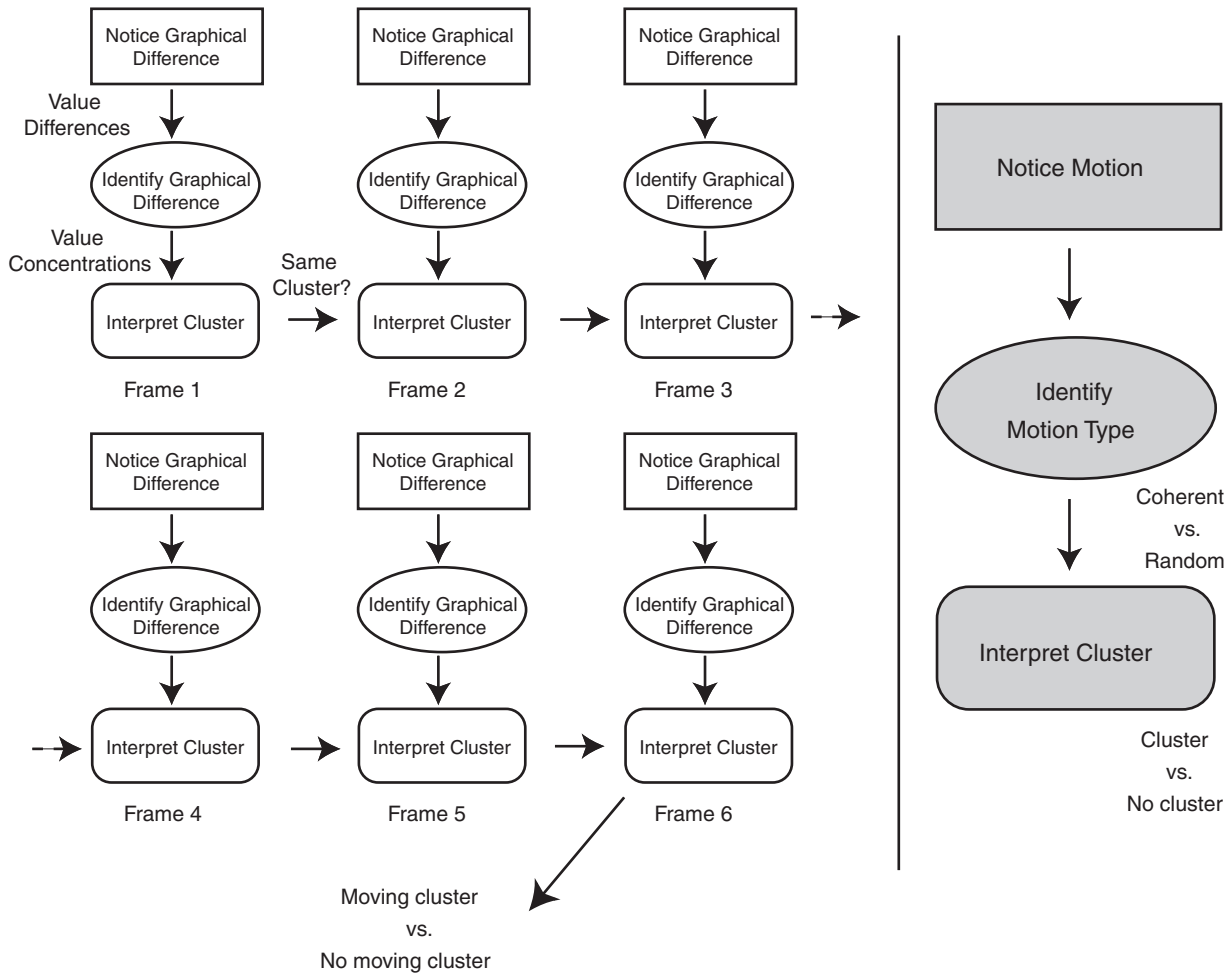


Figure 13. Cognitive demand analysis for the visual identification of moving clusters using static small-multiple (left) and animated maps (right). Animated maps are able to compress the cognitive processes needed to identify moving clusters by focusing on an alternative graphical characteristic of the display (motion instead of color value).

find evidence that this was the case for patterns containing subtle clusters or no clusters. In the animated condition, participants had more success in identifying strong clusters at moderately fast paces. The fastest pace apparently does not allow sufficient time to process each view, but, once processed, a relatively quick transition to the next view seems to help achieve pattern coherence throughout the sequence.

Our results suggest that relative motion may play an important but complex role in the identification of space-time clusters in animation. Several authors have argued that motion is indeed an attention-focusing characteristic of visual displays (Nakayama 1985; MacEachren 1995). Facilitating the perception of motion in animations is therefore a fundamental issue. The perception of motion in animations is an example of what psychologists refer to as the apparent motion problem, in which we perceive motion (that is not directly depicted)

from jumps in position between static small-multiple images. The perception of apparent motion depends on the duration of the stimulus (i.e., the length of time a frame is displayed), the interstimulus interval (i.e., the frame rate), and the distance that an object moves. Generally, if an object moves over a larger distance, it will need a shorter interstimulus interval for the viewer to perceive smooth apparent motion than will an object moving over a shorter distance. If the interstimulus interval is too long, the result is what Wertheimer (1912) described as phi movement, in which the object will appear to move from one location to another without traveling through intermediate locations. In the animated condition, participants correctly identified moving clusters at the lowest rate in trials presented at the fastest pace. This may be due to an inability to process each view at the fastest pace. Participants also had substantial difficulty identifying subtle moving clusters at

the slowest pace. At this pace the combination of less-obvious figure-ground relationships in the subtle clusters, combined with a less-coherent movement of the cluster (i.e., the movement may have appeared stuttered like the phi movement, rather than smooth), may have lessened the appearance of the cells in a cluster having a common fate, thereby increasing the cognitive effort required to identify a cluster. One direction for future research may be to first investigate and describe the relationship between frame rate and the distance moved across the screen that results in a smooth versus a stuttered pattern and then test whether participants are better at identifying clusters from patterns containing each type of movement.

Tversky, Morrison, and Betrancourt's (2002) review examined studies that compared the performance of animated and static small-multiple representations for comprehension tasks, which should require substantially more cognitive effort than our task. In comprehension tasks, we would expect map readers to perform additional cognitive actions, such as *tracking* change in the phenomenon over time, *parsing* the overall map into smaller components for more detailed examination, and then *linking* these components back together to develop an integrated understanding of the phenomenon that the map reader was seeking to understand. Developing this type of understanding also requires tapping into the map reader's conceptual mental structures (i.e., his or her schemata) in order to construct new knowledge. By contrast, the task of visually identifying moving clusters is a task that does not demand that map readers develop an understanding of why they are seeing clusters, but instead requires them to identify whether or not clusters exist (which is certainly a prerequisite step to understanding why they might exist). We suggest that animation was successful for this task because it decreased rather than increased the cognitive load for map readers.

Fabrikant (2005) has recently hypothesized that the reason some map animations studied in the context of comprehension tasks have been unsuccessful is that the perceptually salient features of the maps have not been thematically relevant to developing an understanding of the problem, thereby increasing the cognitive load on map readers who have to "sort through" the perceptual stimuli to find those that are most relevant to the task at hand. Indeed, in a 1999 study of interactive weather map animations, Lowe documented that map readers often extracted perceptually salient information while ignoring that which was thematically relevant. In other words, perhaps map animations have been unsuccessful in some contexts because they have not taken into account the cognitive demands of the task for which they

were intended to be used when they were designed. One possible avenue for further research might be to use Fabrikant's (2005) proposed method of comparing modeled perceptual salience of our map stimuli with eye movement tracking to measure whether map readers are extracting information that is both perceptually salient and thematically relevant (in this case the motion of a cluster).

We did not control for visual acuity when selecting participants for this experiment. Although it is possible that one or more participants had uncorrected vision, we do not believe that it would have had a substantial impact on participants' ability to discern clusters as clusters were approximately an inch in diameter and were viewed from a relatively short distance. Moreover, because we used a within-subjects design, a participant who had difficulty discerning clusters because of uncorrected vision should have had the same problems in both the animated and small-multiple conditions. We suspect that having uncorrected vision would have the same impact on cluster detection as viewing maps with a smaller cell size (and hence cluster radius) would have, but we did not have any test stimuli with smaller cell sizes. However, a future experiment could examine the effect of cluster size on participants' ability to detect moving clusters.

Finally, we did not expect to find a difference in the ability of males and females to correctly identify patterns, as our task was neither related to navigation nor did it require the explicit use of a rotational spatial ability. We only found a gender difference in the animated condition; the difference in pattern identification in the small-multiple condition was not statistically significant. The literature provides two potential explanations for the difference that we found: one social or environmental, and one biological.

One relatively recent socioenvironmental change has been the development and widespread use of videogame technology among both adolescents and adults. Recent research has demonstrated that individuals who regularly play action video games have improved general visual abilities, particularly those related to selective visual attention (C. S. Green and Bavelier 2003). This research also demonstrated that both male and female non-videogame players who played action videogames for one hour a day for ten days showed improved visual attention skills. Media studies research has demonstrated that boys play videogames at almost twice the rate of girls (Woodard and Gridina 2000). Taken together, these results suggest that because many of today's generation of boys practice tasks that develop their visual attention skills to a higher degree than girls, they are

better equipped to identify moving clusters from animations; this hypothesis is one potential explanation for the gender difference we observed. However, because we did not specifically ask participants about their video-game playing practices, we cannot assume that males in our study played videogames regularly or that the females did not. One way in which it would be possible to gather additional evidence in support of this theory would be to run the experiment again, specifically recruiting males and females who do and do not play videogames.

Another potential explanation for the gender difference lies in the biology of our visual system. Sax (2005) has proposed that male eyes may be better hard-wired for sensing motion, while female eyes do better at sensing color and texture because of a difference in the relative number of M-cells (those that sense motion and direction) and P-cells (those that sense color and texture) among males and females. Although this structural difference in the relative amounts of different sensory receptor types has only been conclusively demonstrated in rats (Salyer et al. 2001), similar studies in humans may show the same pattern.

Regardless of what the results of future physiological studies demonstrate, C. S. Green and Bavelier's (2003) study demonstrates that it is possible for both males and females to develop general visual skills with practice. Future cartographic research could investigate whether gender differences are seen in other types of dynamic maps (e.g., interactive animated maps) and for other types of dynamic map-reading tasks, particularly those involving more cognitive processing, such as characterizing change in attributes over time. Additionally, it would be useful to determine the robustness of the gender effect we observed in this relatively small study by conducting other studies that examine gender differences for tasks that are heavily influenced by perceptual factors. Results of such studies could then be applied both to developing training materials to help map readers learn how to best work with dynamic maps and to designing symbolization solutions that may make it easier for less-experienced map readers to work with dynamic maps of any kind.

Acknowledgments

Funding for this work was provided by National Science Foundation Digital Government Grants 9983451, 9983459, and 983461. The authors would like to thank Dr. Mark Harrower for his comments on the initial experimental design; Brian Pacheco, Geoff Hatchard, and Derek Swingle for their assistance with data collection;

the Penn State University students who participated in the research; and the four anonymous referees who provided helpful and constructive critiques of earlier drafts of this article.

Notes

1. Interactivity itself has been shown to increase students' ability to learn, so comparing interactive animations with noninteractive static small-multiple visual representations would confound interactivity and animation. For this reason, Tversky, Morrison, and Betrancourt (2002) decided to exclude studies of interactive animations from their review.
2. Clusters can be considered to be either (1) concentrations of events surrounded by nonevents or (2) areal concentrations within which the mapped phenomenon exists at either lower or higher intensities than surrounding areas.

References

- Acevedo, W., and P. Masuoka. 1997. Time-series animation techniques for visualizing urban growth. *Computers & Geosciences* 23 (4): 423–36.
- Blok, C., B. Kobben, T. Cheng, and A. A. Kuterema. 1999. Visualization of relationships between spatial patterns in time by cartographic animation. *Cartography and Geographic Information Systems* 26 (2): 139–51.
- Carr, D. B., A. R. Olsen, and D. White. 1992. Hexagon mosaic maps for display of univariate and bivariate geographical data. *Cartography and Geographic Information Systems* 19 (4): 228–36, 271.
- Cutler, M. E. 1998. The effects of prior knowledge on children's ability to read static and animated maps. MS thesis, University of South Carolina, Columbia.
- Evans, B. J. 1997. Dynamic display of spatial data-reliability: Does it benefit the map user? *Computers & Geosciences* 23 (4): 409–22.
- Fabrikant, S. I. 2005. Towards an understanding of geovisualization with dynamic displays: Issues and prospects. In: *Proceedings of the 2005 Spring Symposium of the American Association for Artificial Intelligence, Reasoning with mental and external diagrams: Computational modeling and spatial assistance*, Stanford, CA, 21–23 March 2005, ed. T. Barkowsky, C. Freksa, M. Hegarty, and R. Lowe, 6–11. Menlo Park, CA: AAAI Press.
- Gilmartin, P. P., and J. C. Patton. 1984. Comparing the sexes on spatial abilities: Map-use skills. *Annals of the Association of American Geographers* 74: 605–19.
- Green, C. S., and D. Bavelier. 2003. Action video game modifies visual selective attention. *Nature* 423:534–7.
- Green, T. 1989. Cognitive dimensions of notation. In *People and Computers V*, ed. A. Sutcliffe and L. Macauley, 443–60. Cambridge, U.K.: Cambridge University Press.
- Harrower, M. 2002. Visualizing change: Using cartographic animation to explore remotely-sensed data. *Cartographic Perspectives* 39:30–42.
- . 2004. A look at the history and future of animated maps. *Cartographica* 39 (3): 33–42.
- Harrower, M., A. M. MacEachren, and A. L. Griffin. 2000. Developing a geographic visualization tool to support earth

- science learning. *Cartography and Geographic Information Science* 27 (4): 279–93.
- Hegarty, M. 2004. Dynamic visualizations and learning: Getting to the difficult questions. *Learning and Instruction* 14:343–51.
- Hegarty, M., S. Kriz, and C. Cate. 2003. The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction* 21 (4): 325–60.
- Kossoulakou, A., and M.-J. Kraak. 1992. Spatio-temporal maps and cartographic communication. *Cartographic Journal* 29 (2): 101–8.
- Kulldorff, M., and N. Nagarwalla. 1995. Spatial disease clusters—Detection and inference. *Statistics in Medicine* 14 (8): 799–810.
- Lee, P. U., A. Klippel, and H. Tappe. 2003. The effect of motion in graphical user interfaces. *Lecture Notes in Computer Science* 2733:12–21.
- Lloyd, R. E., and R. L. Bunch. 2005. Individual differences in map reading spatial abilities using perceptual and memory processes. *Cartography and Geographic Information Science* 32 (1): 33–46.
- Lowe, R. 1999. Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education* 14:225–44.
- MacEachren, A. M. 1995. *How maps work: Representation, visualization and design*. New York: Guilford.
- Macromedia Flash 5, Macromedia, San Francisco, CA, USA.
- Malinowski, J. C., and W. T. Gillespie. 2001. Individual differences in performance on a large-scale, real-world wayfinding task. *Journal of Environmental Psychology* 21:73–82.
- Montello, D. R., K. L. Lovelace, R. G. Golledge, and C. M. Self. 1999. Sex-related differences and similarities in geographic and environmental spatial abilities. *Annals of the Association of American Geographers* 89: 515–34.
- Nakayama, K. 1985. Biological image motion processing: A review. *Vision Research* 25:625–60.
- Openshaw, S., M. Charlton, and A. Craft. 1988. Searching for leukemia clusters using a geographical analysis machine. *Papers of the Regional Science Association* 64:95–106.
- Peterson, M. P. 1995. *Interactive and animated cartography*. Englewood Cliffs, NJ: Prentice Hall.
- Price, S. 2004. Processing animation: Integrating information from animated diagrams. *Lecture Notes in Artificial Intelligence* 2980:360–4.
- Salzer, D. L., T. D. Lund, D. E. Fleming, E. D. Lephart, and T. L. Hormath. 2001. Sexual dimorphism and aromatase in the rat retina. *Developmental Brain Research* 126 (1): 131–6.
- Sax, L. 2005. *Why gender matters*. New York: Random House.
- Slocum, T. A. 1983. Predicting clusters on graduated circle maps. *American Cartographer* 10:59–72.
- Slocum, T. A., R. S. Sluter, F. C. Kessler, and S. C. Yoder. 2004. A qualitative evaluation of MapTime, a program for exploring spatiotemporal point data. *Cartographica* 39 (3): 43–68.
- Slocum, T. A., S. C. Yoder, F. C. Kessler, and R. S. Sluter. 2000. MapTime: Software for exploring spatiotemporal data associated with point locations. *Cartographica* 37 (1): 15–29.
- Tversky, B., J. B. Morrison, and M. Betrancourt. 2002. Animation: Can it facilitate? *International Journal of Human Computer Studies* 57:247–62.
- Wertheimer, M. 1912. Experimentelle Studien über das Sehen von Bewegung [Experimental studies on motion perception]. *Zeitschrift für Psychologie* 61:161–265.
- Woodard, E. H., and N. Gridina. 2000. Media in the home 2000: The fifth annual survey of parents and children. http://www.annenbergpublicpolicycenter.org/05_media_developing_child/mediasurvey/survey7.pdf (last accessed 23 January 2006).

Correspondence: School of Physical, Environmental and Mathematical Sciences, University of New South Wales-ADFA, Canberra ACT, 2600, Australia, e-mail: a.griffin@adfa.edu.au (Griffin); GeoVISTA Center, Department of Geography, The Pennsylvania State University, 302 Walker Building, University Park, PA 16802, e-mail: maceachren@psu.edu (MacEachren); Department of Geography, University of South Carolina, Callcott Social Sciences Building, 709 Bull Street, Room 127, Columbia, SC 29208, e-mail: hardisty@sc.edu (Hardisty); InfoGraphics Laboratory, University of Oregon, Department of Geography, 1251 University of Oregon, Eugene, Oregon 97403-1251, e-mail: steiner@uoregon.edu (Steiner); Department of Information Technology, Richland County Government, 2020 Hampton Street, Room 3061, P.O. Box 192, Columbia, SC 29202, e-mail: bonan.li@gmail.com (Li).