Change Blindness in Animated Choropleth Maps: An Empirical Study

Carolyn Fish, Kirk P. Goldsberry and Sarah Battersby

Abstract: Animated choropleth maps enable cartographers to visualize time-series data in a way that congruently depicts change over time. However, users have difficulty apprehending information encoded within these displays, and often fail to detect important changes between adjacent scenes. Failures of visual experience, such as change blindness, threaten the effectiveness of dynamic geovisual displays, in which several important changes can occur simultaneously throughout the display. Animated choropleth maps require viewers not only to notice changes but also understand symbolic meanings encoded in rapid transitions between scenes. Graphic interpolation between key frames, also known as "in-betweening" or "tweening", smoothes transitions and lengthens the duration of the transition between scenes in a dynamic sequence. Previous cartographic literature suggests tweening could be a potential solution for change blindness in the cartographic context. This article examines this issue of change blindness in the cartographic context and reports on a human subjects investigation designed to evaluate the influence of cartographic design variables on map readers' change detection abilities. Our results indicate that 1) map readers have difficulty detecting changes in animated choropleth maps, 2) map readers overestimate their own change detection abilities, and 3) tweening influences the legibility of change in animated choropleth maps.

Keywords: Change blindness, choropleth maps, animation

Introduction

Recent developments in technology have enabled cartographers and other designers to represent spatiotemporal processes using dynamic geovisual displays. For example, animated thematic maps are now commonly used to depict spatiotemporal geographic data sets. However, recent research has revealed that viewers often have difficulty apprehending information encoded within dynamic visual displays (Tversky et al. 2002). Some have indicated that the additional attentional burdens of animations may be too much to bear: "our visual attention is typically focused on only one location at a time, so that in a single viewing of an animation, we are likely to miss some of the important component motions (Hegarty et al. 2003: 355)." Similarly, other investigations have revealed that users of animated displays have trouble extracting relevant geographic information relative to users of static displays (Lee et al. 2003), but the specific causes of these failures remain largely unknown. One potential cause may be the surprising inability of observers of animated displays to detect important visual changes that occur during scene transitions; this phenomenon is known as "change blindness" (Grimes 1996; Rensink 2002; Simons and Rensink 2005). However, it remains unknown to what extent change blindness hinders dynamic map reading tasks (Fabrikant et al. 2008; Goldsberry and Battersby 2009). Since several important changes can occur simultaneously throughout the display

Carolyn Fish and Kirk P. Goldsberry, Department of Geography, Michigan State University, East Lansing, MI, USA. Email: <cfish11@msu.edu>, <kg@msu.edu>; **Sarah E. Battersby**, Department of Geography, University of South Carolina, Columbia, SC, 29208, USA. Email: <battersby@sc.edu>.
DOI: http://dx.doi.org/10.1559/15230406384350

during a single transition, animated map displays can potentially overload the cognitive systems of their users (Harrower 2007a; Fabrikant et al. 2008; Goldsberry and Battersby 2009).

Change detection is a key component of contemporary map reading, so it is essential that contemporary cartographic designs accommodate human change detection abilities. This challenge is amplified by the "metacognitive" phenomenon of "change blindness blindness," or viewers' tendency to overestimate their own change detection competencies. In recent non-cartographic studies, participants repeatedly expressed confidence as they failed to detect large changes in visual scenes (Levin et al. 2000 and 2002). Observers not only missed important changes between visual scenes, they were also "blind" to their own change blindness.

Both change blindness and change blindness blindness have specific implications for GIScience, where the inability to apprehend visual changes can result in the misinterpretation of geographic information. Since changes in attributes over time are often symbolized by graphical changes in visual variables (Bertin 1983), the incidence of change blindness in geovisualization commonly involves the failure to notice small visual shifts in position, hue, value, size, or shape. Furthermore, due to change blindness blindness, viewers might not only miss these types of changes, but also incorrectly believe that they interpreted the map correctly. Consequently these observers could be less inclined to review complex display sequences and more inclined to underestimate the overall dynamics of the mapped phenomenon.

This article investigates change legibility in dynamic geovisual displays. More specifically, we report on "change blindness" and "change blindness blindness," and how these perceptual phenomena can reduce the effectiveness of animated choropleth maps. We report the findings of an experiment designed to assess how well people 1) detect shifts in animated choropleth maps, and 2) assess their own change detection abilities. Our research intends to answer three questions:

1. *How well do viewers detect changes in animated choropleth maps?*
2. *Does the cartographic design of scene transitions influence viewers' change detection abilities?*
3. *How does viewers' confidence in their own change detection abilities differ from their accuracy in detecting changes in a cartographic domain?*

Previous Research

Despite the remarkable technological developments that have enabled more complex dynamic geovisual displays, an important set of new challenges has emerged. Since cost and access are less of an obstacle, contemporary research questions have less to do with computation and more to do with the design, perception, and cognition of dynamic geovisual displays. When it comes to animated maps, the hurdle is no longer technological - it is the limited visual and cognitive processing capabilities of the map reader (Harrower 2007a).

For a viewer to completely comprehend the meaning encoded in a dynamic map, the viewer not only has to *notice* all of the graphical changes between scenes, but they also must *understand the meaning encoded within the changes*. This requires readers to first perceive and understand the original state of the geographic information, then notice the changes to the display, attend to those changes, and decode the meaning of the transformation. Goldsberry and Battersby (2009) identified three levels of change detection applicable to dynamic thematic map reading:

- *Change Detection Level 1*: the reader notices the presence or absence of a change;
- *Change Detection Level 2*: the reader detects whether the change was an increase or decrease in the mapped phenomena (only in quantitative thematic maps);
- *Change Detection Level 3*: the reader fully detects the change and makes the connection between the origin and destination states to understand the meaning of the cartographic transition.

Beck and Levin (2003) tested this idea in the psychology discipline and found that participants were able to correctly identify changes (Change Detection Level 1), however, they were unable to accurately recall the origin state of the visual before the change (Change Detection Level 3). In map reading Change Detection Level 3 is vital to complete comprehension of an animated map. To achieve the highest level of change detection and to fully understand the meaning of the dynamic geovisual transformation (Level 3), map designs need to afford readers time not only to notice the origin and destination states, but also perceive the transition between the two.

As a result of conventional design strategies, many important change signals in dynamic maps are only implicitly presented to observers. Changes in geographic information are often symbolized by sudden graphical changes in the visual variables, but the intrinsic meaning of the underlying change is never explicitly displayed. To apprehend this intrinsic meaning a reader must "read between the lines"; first, by recalling the original state of a map feature, second, by noticing its new state, and third, by deciphering the symbolic difference between the two.

Previous researchers have suggested that by graphically interpolating between scenes (i.e. tweening), cartographers may be able to make cartographic changes more legible (Fabrikant et al. 2008) by extending the amount of display time dedicated to transitions. Tweened transitions may afford readers more time to identify, and attend to ongoing changes between adjacent scenes. Simons et al (2000), however, found the opposite, that smooth transitions often masked the change and prevented animation viewers from noticing even drastic changes. Their experiment, however, used a 12-second tween that could be argued as too gradual a transition to elicit a change signal. These ideas remain untested, however, in the cartographic domain, thus the extent to which gradual transitions facilitate or impede change detection in animated maps remains largely unknown.

Methods

An experiment was designed to investigate 1) the legibility of change in animated choropleth maps, 2) the influence of tweening on change detection and 3) the prevalence of change blindness blindness in a cartographic context. Seventy-eight students, mostly undergraduates, from 46 different majors at Michigan State University participated in the study. Each participant was compensated $10 for about 30 minutes of his or her time.

Experimental Design

After a short pretest, participants in the study completed 108 total questions. Every question involved a two-scene animated choropleth map sequence; the first scene would appear on the display, and then the map would transition to the second scene. After the transition completed, a red rectangle would appear to highlight one of the 64 enumeration units of the map (Figure 1). The number of changes on the map that occurred between scenes ranged from 11 to 50 of the total 64 units. The mean number of units that changes between scenes was 22.44. In the test animations, the second scene remained visible on the screen for an unlimited amount of time until the participant answered the question. The test consisted of 108 total questions, which fell into two categories. Question Type #1 is shown in Figure 2. These questions required participants to simply indicate whether he or she believed the highlighted unit's symbology had changed during the transition. These questions

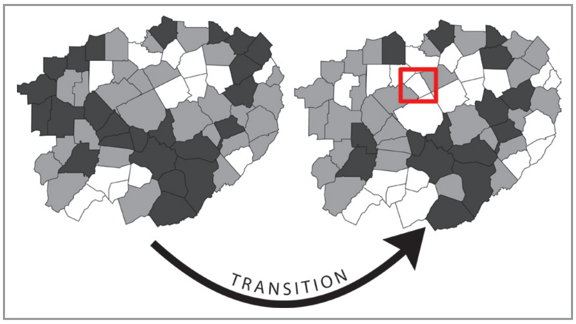

Figure 1. Example of a transition between two scenes. The red rectangle was intended to highlight an enumeration unit of the map. The participant was required to determine whether the highlighted unit changed in lightness between the two scenes.

Figure 2. Example of Question Type #1. In this question, participants simply had to indicate whether they had noticed a change or not. This question was designed to understand how well participants were able to detect changes at Level 1.

were designed to assess accuracy for Change Detection Level 1, the basic ability to notice the presence or absence of cartographic changes.

In Question Type #2, shown in Figure 3, participants were asked to recall the original symbology of the highlighted unit prior to the scene transition. To correctly respond to these questions participants would not only have to notice the presence or absence of a change, but must also be able to recall the origin state of the map feature prior to a change. This level of change detection is necessary to fully apprehend the cartographic meaning encoded in the display. In other words, simply noticing a shifting or persisting behavior (Question Type #1) does not mean the map reader fully comprehends the transition that occurred between scenes in the animated map sequence.

Figure 3. Example of Question Type #2. In this question, participants were instructed to indicate the shade of highlighted unit before the scene change. This question was designed to understand how well participants were able to detect changes at Level 3. Participants not only had to grasp whether the unit had changed, but also the origin state of the unit before the change.

In this experiment, change detection was assessed for three different transition design conditions:

1. *Abrupt (Figure 4)*
2. *Delayed Smooth (Figure 5)*
3. *Continuous Smooth (Figure 6)*

A "within-subjects" design was used; each subject was exposed to all three conditions. In all three conditions participants saw a two-second dynamic choropleth map sequence consisting of an origin state and a destination state; however, the transition between scenesvaried across conditions.

In the "Abrupt Condition", the origin state (map scene 1) was shown for a full two seconds then abruptly transitioned to the destination state of the map (map scene 2), thus completing the sequence without any tweening (Figure 4). A 250-millisecond delay was inserted between the end of the transition sequence and the onset of the highlight rectangle in the Abrupt Condition to prevent distraction (O'Regan et al. 1999). There were 36 questions about maps in the Abrupt Condition, divided equally between the two question types.

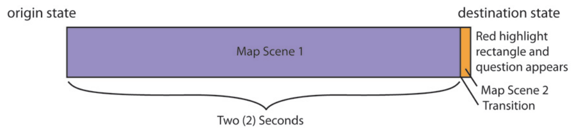

Figure 4. The Abrupt Condition. In this condition, the static first scene (purple) was shown for two (2) seconds, followed by the second scene (orange) which was shown for 250 milliseconds before the highlight rectangle and question appeared.

In the "Delayed Smooth Condition", the origin state was shown for one second before a one-second tweened transition occurred (Figure 5). Upon the completion of the transition, the highlight rectangle and the question appeared. Once again, participants had an unlimited amount of time to view the second scene and answer the question. Thirty-six questions were asked about maps within the Delayed Smooth Condition, 18 of Question Type #1, and 18 of Question Type #2.

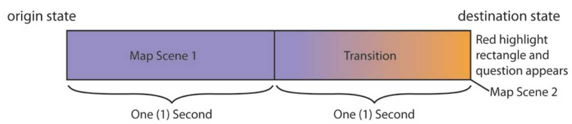

Figure 5. The Delayed Smooth Condition. In this condition, the static first scene (purple) was shown for one (1) second, followed by a one (1) second transition between the first and second scene. Once the transition ended, the highlight rectangle and question appeared.

In the "Continuous Smooth Condition" the display was in a perpetual state of tweening from the onset of origin state through the completion of the transition to the destination state (Figure 6). This condition most closely adheres to the congruence principle of graphics by mimicking the continual passage of time with a perpetual graphical transition (Tversky et al. 2002). Immediately following the transition, the red highlight rectangle and the question appeared. Thirty-six questions were asked about maps within the Continuous Smooth Condition divided equally between the two question types.

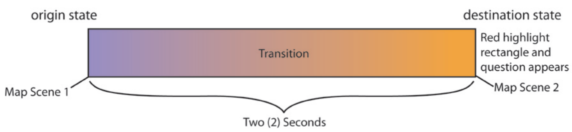

Figure 6. The Continuous Smooth Condition. In this condition, no static scene was shown, the map immediately began to transition upon appearance on the screen and continued to transition for two (2) seconds. Once the transition ended, the highlight rectangle and question appeared.

Animated descriptions of each of the question and condition types can be found at: https://www.msu.edu/~kg/example.swf

Table 1. Illustration of the number of questions of each of the two question types and three condition types each of the participants saw.

Breakdown of Questions and Condition Types				
	Abrupt	*Delayed Smooth*	*Continuous Smooth*	*Total*
Questions Type 1 (CD Level 1)	18	18	18	54
Questions Type 2 (CD Level 1)	18	18	18	54
Total	36	36	36	108

Table 1 illustrates the breakdown of the test:

Following every question in the study, participants were asked to indicate their confidence on the previous question by indicating "confident", "neutral", or "unsure" (Figure 7). There were a total of 108 confidence questions.

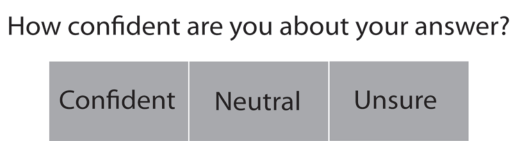

Figure 7. Example of confidence question. Participants were able to indicate that they were "confident", "neutral", or "unsure" of their answer on the previous question. There were 108 confidence questions, one following each of the change detection questions.

Materials
The study was conducted on Dell Precision 690 Computers with Dell 1907 FPt computer displays with a screen resolution of 1280 x 1024, and a refresh rate of 60 Hertz. The brightness was set to 100 and the contrast was set to 50 on all of the displays. The questions in the study were designed in Adobe Flash CS3 (Adobe Systems, Inc. 2007) using the default frame rate of 12 frames per second. The experiment was administered using Mozilla Firefox (Mozilla Corporation 2009) in "full-screen" mode.

Procedure
Each participant completed the experiment in a Windows computer lab. During the study, subjects were asked to enter the testing room, were assigned computers, and were thanked for their participation. The computers screens were set 30 centimeters from the edge of the desk; however, participants were allowed to sit at whatever distance from the screen they felt was comfortable. Before starting the actual test questions, each participant signed a consent form, answered basic demographic and familiarization questions, and took part in a pretest session on the computer. In the pretest, the participants were shown four practice questions and were given general information about the study.

Upon completion of the test, participants took part in a posttest. The posttest consisted of open-ended opinion questions about their feelings towards the test, the transitions, and the training. The goals of the posttest were to 1) assess whether participants had different opinions about the different types of transitions they saw, and 2) ascertain whether the training was adequate at familiarizing the participants with the types of questions they would see. Once all of the participants in the room had finished the test, the participants were given their $10 compensation and departed.

Results

Signal Detection Theory
Signal Detection Theory (SDT) is commonly used as a means to determine the strength of a signal within a dynamic stimulus, but has rarely been used to assess signals for geographic infor-

mation tasks (Griffin and Bell 2009). SDT allows us to quantify map readers abilities to discriminate changes and biases. For "yes/no" questions decision responses can be categorized into four types of potential SDT outcomes (Table 2): Hits, Correct Rejections, False Alarms, and Misses (Macmillian and Creelman 1991). Each of these outcomes is defined below in an animated map context:

Hits: the participant correctly identifies that a mapped unit has changed during a scene transition.

Correct Rejections: the participant correctly identifies that a map unit has not changed during a scene transition.

False Alarms: the participant incorrectly indicates that a map unit has changed, when in fact the mapped unit did not change.

Misses: the participant incorrectly indicates that a map unit has not changed during a map transition, when in fact the mapped unit changed during a scene transition.

The four possible outcomes are shown in (Table 2).

Table 2. The SDT outcome array. Categorization of map readers' potential decisions based on what they were shown.

	Participant Decision	
Reality	*Change*	*No-Change*
Change	Hit (H)	Miss (M)
No-Change	False Alarm (FA)	Correct Rejection (CR)

Each of these four outcomes reveals something important; cartographers should seek to maximize Hits and Correct Rejections while minimizing the incidence of False Alarms and Misses. SDT was used to measure discriminability of a change and participant bias to indicate either "yes" or "no" in this experiment. The measure of discriminability (d') is defined for two choice questions (Question Type #1) as:

$$d' = z(H) - z(FA)$$

As d' increases the discriminability of a change increases, meaning participants had an easier time distinguishing cartographic change behaviors from persisting behaviors.

Due to the nature of forced answer trials, when participants are unsure of an answer, they must guess; depending on the nature of the stimulus, participants may adopt guessing strategies that can skew test results. There are several ways to measure these guessing tendencies, also known as participant bias. In this study the formula used was:

$$\text{criterion } (c) = \tfrac{1}{2}\,(z(H) + z(FA))$$

A positive c indicates participants had a tendency to say "no change" while a negative c indicates a tendency to say "change". A zero c indicates no bias toward either "no change" or "change".

Detection Accuracy

Change detection performance was evaluated for all 108 questions of the study. Participants were evaluated on their abilities to detect changes as well as persisting behaviors (i.e. the absence of a change) between map scenes. Table 3 summarizes the overall test performance, across all conditions and question types for the entire cohort.

Table 3. Participant performance for all 108 questions of the experimental test. This table is the combined results of both question types and all three conditions. In the experimental test there were twice as many questions where the highlighted unit had changed (72 of the 108 questions), making the correct answer a Hit, than when the highlighted unit had persisted through the transition (36 of the 108 questions).

Total Performance		
	Change (Hit Rate)	*No Change (Correct rejection rate)*
Number of Questions	72	36
Mean	0.454	0.656
Standard Deviation	0.098	0.133
Minimum	0.236	0.167
Maximum	0.708	0.889

Change Detection Level Performance

Participants were also evaluated on their abilities to detect changes at two different levels of change detection. Level 1 Change Detection refers to questions where participants were asked simply to notice the presence or absence of change (Question Type #1), while Level 3 Change Detection refers to questions where participants were required to recall the origin state of the highlighted unit (Question Type #2). Participants had higher correct answer rates (Hits plus Correct Rejections) for Change Detection Level 1 questions (m = 34.49 / 54) than Change Detection Level 3 (m = 22.77 / 54).

Accuracy by Condition Type- Change Detection Level 1

To examine the effect of different transition designs on change legibility, participants were evaluated on their ability to detect changes with three different transition types: Abrupt, Delayed Smooth, and Continuous Smooth. Table 4 illustrates participant performance on Change Detection Level 1 questions across conditions.

Table 4. Participant performance for Change Detection Level 1 across conditions. The high d' of the Continuous Smooth Condition indicates that it is easier to discriminate

Performance for Change Detection Level 1 Across Conditions			
	Abrupt	*Delayed Smooth*	*Continuous Smooth*
Hit Rate	0.627	0.647	0.596
False Alarm Rate	0.545	0.348	0.258
Correct Rejection Rate	0.455	0.652	0.742
Miss Rate	0.373	0.353	0.404
d'	0.206	0.771	0.897
criterion c	-0.239	-0.001	0.195

Hit Rates were similar across conditions, while False Alarm Rates varied considerably. The highest Hit Rate was in the Delayed Smooth Condition, while the lowest False Alarm Rate was in the Continuous Smooth Condition. The differences in the d' measure across design conditions suggest that participants were better able to discern changes within the two tweened conditions than they were in the Abrupt Condition. The highest d' in the Continuous Smooth Condition indicates that longer, more gradual transitions enable readers to better discern differences between cartographic change and persistence. Conversely, the lower d' in the Abrupt Condition indicates that participants had more difficulty discerning change behaviors from persisting behaviors when tweening was not applied. This lower d' value is primarily due to the high rate of False Alarms in this condition. A one-way repeated measures ANOVA indicated that there was a significant difference in d-prime values within subjects across conditions, $F(2,234 = 30.34$, $p<0.0001$. A comparison of the main effect using a Bonferroni corrected paired comparison indicated that were significant differences between the Abrupt Condition and the two other conditions, $p<0.0001$ for both; however, the Delayed Smooth Condition and the Continuous Smooth Condition were not significantly different from one another, $p= 0.732$.

The criterion values, c reveal additional important differences across conditions; the negative value in the Abrupt Condition suggests that participants tended to think the stimulus was more volatile than it actually was, and consequently were biased to picking "change" over "no change" – the exact opposite phenomena occurs in the continuous condition, which may suggest changes in this condition were not as perceptually salient. In the Delayed Smooth Condition, this resulted in a c very close to zero, indicating that participants had very little bias.

Participants' performance was considerably lower for Question Type #2 (Change Detection Level 3), in which they were asked not only to detect the presence or absence of change, but also to characterize the nature of the change. We adapted the original four possible outcomes used for the two-alternative questions (Question Type #1), to accommodate the three-alternatives of Question Type #2. The new categorization divides Hits into "Hits" and "False Hits;" a Hit is when the participant is able to correctly answer the question by indicating the original symbology of the highlighted unit. A False Hit is when the participant correctly identified that the unit had changed, but incorrectly indicated the origin state of the unit before the change.

For Question Type #2 participants had the highest Hit Rates (0.370) and False Hit Rates (0.243) in the Abrupt Condition. However, when there was a persistence behavior, participants also produced the highest rate of False Alarms (0.370) and fewest Correct Rejections in this condition. In the Delayed Smooth Condition, participants had a lower rate of Hits (0.283) than in the Abrupt Condition. However, during a persistence behavior, participants produced fewer False Alarms (0.286) in this condition than they did in the Abrupt Condition indicating that they were less biased to believe there had been a change and had a higher Correct Rejection Rate (0.714). Finally, in the Continuous Smooth Condition, participants produced the lowest rate of Hits (0.201), but the False Alarm Rate (0.256) was also the lowest (Table 5).

Table 5. Participant performance for Change Detection Level 3 across conditions. Hit Rates, False Hit Rates, and Miss Rates refer to possible outcomes where there is a change behavior. Correct Rejections and False Alarms are possible when there is a persistence behavior. A true Hit is when the participant is able to correctly answer a Change Detection Level 3 question by correctly identifying the original state of the highlighted unit, while a False Hit Rate is when the participant incorrectly responds to the Change Detection Level 3 question by selecting the incorrect origin state, but was able to identify that the unit did change.

Performance for Change Detection Level 3 Across Conditions			
	Abrupt	*Delayed Smooth*	*Continuous Smooth*
Change Behavior			
Hit Rate	0.370	0.283	0.201
False Hit Rate	0.243	0.242	0.232
Miss Rate	0.387	0.475	0.567
Persistence Behavior			
Correct Rejection Rate	0.630	0.714	0.744
False Alarm Rate	0.370	0.286	0.256

Detection Accuracy by Change Type

Changes within a dynamic choropleth map can be effectively categorized using cross-change characterization arrays (Monmonier 1975; Goldsberry 2004). In the array, the diagonal elements from upper left to lower right indicate the units that did not change while the off-diagonal elements indicate the units that did change and the nature of those changes. The maps used in this experiment were three class maps with “high”, “medium”, and “low” categories. A three-class map sequence can exhibit nine possible transition behaviors. We summarize the cohort’s overall performance across each transition behavior below.

The cohort’s performance was strongest for Correct Rejections, when participants correctly identified a persisting behavior between map scenes (Figure 8). Participants were less successful at identifying shifting behaviors, and least successful when the unit shifted from “high” to “low” with only a 40% accuracy rate. Hit Rates, where participants correctly identified there was a change, however, were relatively similar for all types of change, ranging from 40% to 56%.

Confidence Results

Following each question, participants were asked to indicate whether they felt “confident,” “neutral,” or “unsure” about their answer. Following a correct answer, participants indicated that they were “confident” 54% of the time, “neutral” 32% of the time and “unsure” only 14% of the time. Following incorrect responses, the mode response was “neutral” (41%) however participants indicated “confident” (38%) more than “unsure” (21%). The percentage of correct answers did significantly differ by confidence rating, $\chi^2(2, N = 8424) = 215.78, p < 0.0001$.

Further evaluation of the confidence ratings by change detection level reveal that participants were less confident on Change Detection Level 3 questions. However, despite the lower confidence ratings for Change Detection Level 3, “unsure” was still the least popular answer.

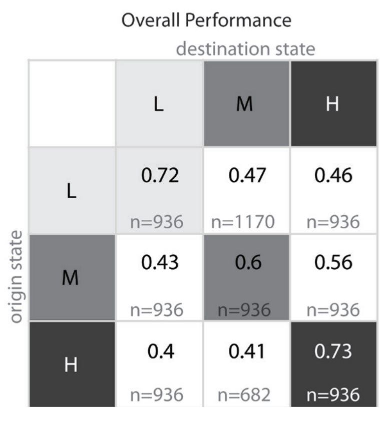

Figure 8. Cross-classification array of accuracy rates for all 108 questions in the experiment. These rates indicate how well participants performed on questions for each of the nine possible transitions. The results indicate that the highest rates of detection occurred when the unit remained the same, the diagonals from top left to bottom right. For example, participants were 73% accurate on detecting the absence of a change when the unit remained “high” between the two scenes.

Discussion

How well do map readers detect changes in dynamic choropleth maps?

The results from this experiment reveal that change blindness in the cartographic context, the failure to detect change between scenes in a dynamic map scene, can undermine the effectiveness of cartographic animation. This problem is further complicated by users' tendency to overestimate their own change detection abilities, meaning they fail to realize that they missed many cartographic signals, which could potentially cause misinterpretations of larger spatiotemporal trends. Lastly, our results indicate that map readers can identify change at low levels (simply noticing the presence or absence of cartographic shifts) more readily than they can at higher levels (the ability to recall the original state of a transitioning map element).

Maps graphically depict geographic information, and cartographers, unlike other designers, cannot fully control the look and behavior of geovisual displays. Consequently, many changes commonly occur across a display at a single moment in time. As suggested by previous researchers our visual attention is typically focused on one location at a time, so it may be impossible to apprehend dozens of simultaneous change signals distributed widely across a display (Hegarty et al. 2003). Furthermore, if the primary utility of these dynamic displays is their ability to portray changes over time and space (Harrower 2007a), the potential implications of change blindness threaten to undermine the effectiveness of dynamic geovisualization. When users miss changes they fail to apprehend the meaning encoded within the display. Given the inherent complexity of these displays it is requisite for readers to understand symbolic meanings, not just notice change. For example, in a dynamic choropleth map of unemployment changes over time, the consequences of failing to detect a change from the lowest class to the highest class induces a failure to understand the dynamics in the underlying unemployment information. Unlike movies or other stimuli used in previous change blindness studies, simply noticing cartographic changes is only the beginning of many geovisual tasks.

The results of this study also demonstrate that lower levels of change detection are easier to achieve than higher levels (Goldsberry and Battersby 2009). To comprehend the meaning of a dynamic choropleth map, the reader must not only notice a change (Change Detection Level 1) in a particular unit, but also must understand how that unit changed over time (Change Detection Level 3). Although a user may be able to deduce that if a map symbol has shifted, it must have changed from a different symbol, the full meaning of the transition is lost if the reader is not able to recall the origin state of that symbol. While complete comprehension of animated choropleth maps requires the user to achieve Change Detection Level 3, this is not to reduce the importance of detecting changes at lower levels. Lower levels of change detection are essential to understanding these types of maps because detection of a simple change can lead to further user interaction that in turn leads to more complete comprehension.

Finally, analysis of cross-classification change arrays reveals that different types of shifting elicit different change detection rates. The results suggest that choropleth map readers are more adept at detecting persistence (Correct Rejections) behavior than they are at detecting changes (Hits). These results also reveal that participants had a bias to believe that units transitioned from the neighboring class as opposed to a more distant class. For example, when an enumeration unit transitioned from the "low" to the "high" class, participants frequently indicated that they believed the unit shifted from the "middle" class. This type of error may reveal a prior notion that more drastic change may seem unlikely. More research is needed to understand how user biases can affect the comprehension of dynamic geovisual displays.

Does transitional design matter?

The results indicate that transitional designs do influence change detection accuracy. Participants performed better within the tweened conditions. While test performance was similar across design conditions, *d'* and *c* indicated that it was easier for participants to distinguish a change and were less biased to assume there was a change in the both tweened conditions; while the Abrupt Condition results had marginally higher Hit Rates, False Alarms were much more prevalent and participants had problematic change biases.

One reason that participants may have been vulnerable to False Alarms in the Abrupt Condition may be related to the general "loudness" of change signals inherent in abrupt changes. Because all of the change occurs suddenly, in adjacent animation frames, changes are instantaneous and many map elements across diverse portions of the display become lighter or darker simultaneously. Essentially, changes within the Abrupt Condition are noisier; they are very salient perceptually, but also so instantaneous that readers do not have sufficient time to attend to the specific map elements that are shifting. This elevates the reader's bias to believe changes have occurred. The Abrupt Condition seems to induce Harrower's (2007b) idea of animated simultaneous contrast, the phenomenon in which readers falsely perceive changes in persisting elements due to noisy change behavior of surrounding elements.

The results also expose issues associated with gradual transitions that may be related to Rensink's (2002) two types of change detection: completed and dynamic. The Abrupt Condition represents an example of completed change detection, where change events are sudden and immediate. Dynamic change detection, conversely, occurs when the viewer notices a transition event during its development. For example, in this experiment's smoothly changing conditions, it was possible for the participant to notice and attend to a transition after its onset, but prior to its conclusion. These sudden changes may cause map readers to become confused when watching the smoothly changing transitions. For example, when a unit transitioned from the "low" class to the "high" class, it was possible that a map reader may have noticed the change during the midpoint of the transition (while the unit was in the "middle" class) and thought that the unit changed from the "middle" class to the "high" class instead from the "low" class.

Confidence and Change Blindness Blindness

Finally, this research attempted to understand to what extent animated map readers experience change blindness blindness, the overestimation of one's own change detection abilities. Consistent with findings in non-cartographic perceptual research (Levin et al. 2000; 2002), participants in this study were generally overconfident, but not very accurate. One participant in this study even noted that the task "was very easy." Further, participants rarely choose the "unsure" choice when rating their confidence, despite the overall failure to detect changes in many of the maps. This result indicates that map readers were highly confident but not highly accurate when detecting changes in animated choropleth maps.

This result is consistent with other change blindness blindness research. Levin and his colleagues (2000 and 2002) found that viewers believed they would see a change, even when 100% of viewers failed to notice the change. In Levin's experiments, participants were asked to indicate whether they believed they would see a change before they saw it. In this study, participants were asked to rate their confidence in change detection after each change detection question. However, even when asked after failing to detect the change, participants were confident they saw it correctly.

The tendency to select "confident" suggests another problem that threatens the effectiveness of these types of displays; if users commonly miss changes – yet are confident they "saw" everything - they are likely to underestimate the overall dynamics of the geographic information represented in these types of maps. This is akin to other perceptual issues that hinder map comprehension such as Flannery's (1971) seminal finding that map readers commonly underestimate the geometric sizes of large circles on proportional symbol maps. More research is needed to explore how cartographers can mitigate this problem.

Previous research has also argued that interactivity helps map readers deal with the added complexities of dynamic geovisual displays. By enabling users to replay, pause or toggle between dynamic sequences, designers offer users more chances to review events they may have missed within the sequence. However, the effects of change blindness blindness may cause map readers to falsely believe they have correctly perceived more displays than they actually have, and thus map readers may underutilize interactivity functions. Basically, map readers who are confident they already "saw" what happened are not going to use a replay button.

Limitations

This study examined map readers' abilities to detect the presence or absence of change in a dynamic choropleth map sequence. Although the findings shed light on emerging perceptual issues that threaten the effectiveness of animated choropleth maps, it is clear there are many important limitations:

- *Static Scene Duration*. As a means to ensure the overall duration of each condition was identical, the display duration of the origin state (the first scene in the sequence) varied by condition. The Abrupt Condition had the longest static scene duration (two seconds), followed by the Delayed Smooth Condition (one second), and finally the Continuous Smooth Condition (approximately 100 milliseconds). According to previous research, the amount of time given to encode information is directly related to the amount of information the observer is able to retain after the change took place (Hollingworth et al. 2001). These conditional design differences were designed to be conceptually identical to realistic design variables that cartographers employ in these types of animated maps; however, it is possible that these differences could contribute to the observed accuracy differences across conditions.
- *Magnitude of Change and Map Complexity*. Several participants noted that it was difficult to detect changes with the high number of enumeration units in the map (n=64), as well as the large magnitude of change that occurred between many of the map scenes. In this experiment, the lowest magnitude of change value was 11 units – meaning 11 of the 64 units shifted during the transition - and the highest magnitude of change was 50. These high magnitudes of change are a result of the underlying data, yearly average unemployment rates by county between 1990 and 2008. There is a possibility that the high magnitude of change of the underlying information may have influenced the cohort's success, but this remains unknown.
- *Congruent Data Representations*. Battersby and Goldsberry (2010) argue that classed maps should not be tweened because it implies there is a smooth transition between the two discrete classes, when in fact there is commonly no numerical gap between two bins of a classed choropleth map, thus graphic interpolation implying otherwise is inappropriate. This study used tweening for a classed map, which may have caused confusion among some participants.
- *Map Tasks*. It should be noted that this study only tested one particular map task; participants were required to notice change in one enumeration unit during a singular complex transition. Other common map tasks not tested here include evaluating 1) one unit over many transitions, 2) many units over one transition, or 3) many units over many transitions. It is important to note that change detection rates may differ as the detection tasks differ. Future research should examine change detection abilities in other types of task domains as well as in other types of dynamic geovisual displays. In this case, the investigation examined the atomic level of change detection in animated choropleth maps, the ability to notice changes in a singular display element across a singular transition.
- *The Splat Contingent of the Highlight Rectangle*. The visual onset of the red rectangle shortly after the end of the transition in this study may induce change blindness by presenting additional changes to the already dynamic map scene. In the Abrupt Condition, a delay of 250 milliseconds was added between the onset of the second map scene and the onset of the rectangle. However, this delay may not be enough time to avoid the "splat contingent" of change blindness (O'Regan et al. 1999).
- *Color*. This study tested animated choropleth maps in gray scale. However, it has been suggested that color may help reduce change blindness in these types of displays. Future experiments similar to the one presented here may benefit from the use of color in the map.

Conclusion

This experiment was designed to investigate and describe when map readers fail to notice, attend to, and perceive seemingly obvious changes in animated choropleth maps. Our intention was to examine how the more general perceptual phenomena of change blindness, and change blindness blindness affect animated map reading. The results indicate that map readers do experience change blindness and thus regularly fail to detect basic changes within animated choropleth maps. In turn, this phenomenon hinders the comprehension of the meanings encoded in these types of maps. Furthermore, despite missing many of the changes within a display, observers commonly believe they identified cartographic changes that they in fact did not.

Dynamic geovisual displays, specifically animated choropleth maps, have the potential to help people understand spatiotemporal geographic information, however, as these displays become more and more complex, change blindness and change blindness blindness each present important new challenges for the map reader. More research is needed to better understand these issues and to inform future designs that can overcome or minimize their effects.

ACKNOWLEDGEMENTS

Funding for this research wasprovided by the Michigan State University Department of Geography Graduate Office Fellowship and the Masters Thesis Grant from the Cartography Specialty Group of the Association of American Geographers.

REFERENCES

Adobe Flash Cs3 Version Creative Suite 3. Adobe Systems, Inc., San Jose, California.

Battersby, S. and K. Goldsberry. 2010. Considerations in Design of Transition Behaviors for Dynamic Thematic Maps. *Cartographic Perspectives,* 65.

Beck, M.R., and D.T. Levin. 2003. The Role of Representational Volatility in Recognizing Pre-and Postchange Objects. *Perception & Psychophysics* 65(3): 458.

Bertin, J. *Semiology of Graphics*. University of Wisconsin Press, 1983.

Fabrikant, S.I., S. Rebich-Hespanha, N. Andrienko, G. Andrienko and D.R. Montello. 2008. Novel Method to Measure Inference Affordance in Static Small-Multiple Map Displays Representing Dynamic Processes. *The Cartographic Journal* 45(3): 201-215.

Flannery, J.J. 1971.The Relative Effectiveness of Some Common Graduated Point Symbols in the Presentation of Quantitative Data. *The Canadian Cartographer* 8(2): 96-109.

Goldsberry, K. 2004. *Stabilizing Rate of Change in Thematic Map Animations*. Unpublished Masters Thesis. University of California Santa Barbara.

Goldsberry, K. and S. Battersby. 2009. Issues of Change Detection in Animated Choropleth Maps. *Cartographica: The International Journal for Geographic Information and Geovisualization* 44(3): 201-15.

Griffin, A.L. and S. Bell. 2009. Applications of Signal Detection Theory to Geographic Information Science. *Cartographica: The International Journal for Geographic Information and Geovisualization* 44(3): 145-58.

Grimes, J. 1996. Scenes across Saccades. *Perception* 5: 89-110.

Harrower, M. 2007a. The Cognitive Limits of Animated Maps. *Cartographica: The International Journal for Geographic Information and Geovisualization* 42(4): 349-57.

Harrower, M. 2007b. Unclassed Animated Choropleth Maps. *The Cartographic Journal* 44(4): 313-20.

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): 209-49.

Hollingworth, A., C.C. Williams and J.M. Henderson. 2001. To See and Remember: Visually Specific Information Is Retained in Memory from Previously Attended Objects in Natural Scenes. *Psychonomic Bulletin & Review* 8(4): 761.

Lee, P., A. Klippel and H. Tappe. 2003. The Effect of Motion in Graphical User Interfaces. *Lecture Notes in Computer Science* 2733: 277-90.

Levin, D.T., S.B. Drivdahl, N. Momen and M.R. Beck. False Predictions About the Detectability of Visual Changes: The Role of Beliefs About Attention, Memory, and the Continuity of Attended Objects in Causing Change Blindness Blindness. *Consciousness and Cognition* 11(4): 507-27.

Levin, D.T., N. Momen, S.B. Drivdahl and D.J. Simons. 2000. Change Blindness Blindness: The Metacognitive Error of Overestimating Change-Detection Ability. *Visual Cognition* 7(1): 397-412.

Levin, D.T. and D.J. Simons. 1997. Brief Reports Failure to Detect Changes to Attended Objects in Motion Pictures. *Psychonomic Bulletin & Review* 4(4): 501-06.

Macmillan, N.A. and C.D. Creelman. 1991. *Detection Theory: A User's Guide*. Cambridge

University Press Cambridge.

Monmonier, M.S. Class Intervals to Enhance the Visual Correlation of Choroplethic Maps. *Cartographica: The International Journal for Geographic Information and Geovisualization* 12(2): 161-78.

Mozilla Firefox. Mozilla Corporation.

O'Regan, J.K., R.A. Rensink and J.J. Clark. 1999. Change-Blindness as a Result of Mudsplashes. *Nature* 398(6722): 34.

Rensink, R.A. 2002 Change Detection. *Annual Review of Psychology* 53(1): 245-77.

Simons, D. J., S.L. Franconeri and R.L. Reimer. 2000. Change Blindness in the Absence of a Visual Disruption. *Perception* 29: 1143-54.

Simons, D.J. and R.A. Rensink. 2005. Change Blindness: Past, Present, and Future. *Trends in Cognitive Sciences* 9(1): 16-20.

Tversky, B., J.B. Morrison and M. Betrancourt. 2002. Animation: Can It Facilitate? *International Journal of Human Computer Studies* 57: 247-62.