# Effectiveness of Animation in Trend Visualization

George Robertson, Roland Fernandez, Danyel Fisher, Bongshin Lee, and John Stasko

**Abstract**— Animation has been used to show trends in multi-dimensional data. This technique has recently gained new prominence for presentations, most notably with Gapminder Trendalyzer. In Trendalyzer, animation together with interesting data and an engaging presenter helps the audience understand the results of an analysis of the data. It is less clear whether trend animation is effective for analysis. This paper proposes two alternative trend visualizations that use static depictions of trends: one which shows traces of all trends overlaid simultaneously in one display and a second that uses a small multiples display to show the trend traces side-by-side. The paper evaluates the three visualizations for both analysis and presentation. Results indicate that trend animation can be challenging to use even for presentations; while it is the fastest technique for presentation and participants find it enjoyable and exciting, it does lead to many participant errors. Animation is the least effective form for analysis; both static depictions of trends are significantly faster than animation, and the small multiples display is more accurate.

**Index Terms**—Information visualization, animation, trends, design, experiment.

---◆---

## 1 INTRODUCTION: TREND VISUALIZATION

Informally, the term *trend* means to have a general tendency (Webster's Dictionary). A trend in data is an observed general tendency. The most common way to see a trend in data is to plot a variable's change over time on a line chart or bar chart. If there is a general increase or decrease over time, this is perceived as a trend up or down. If there is a general increase/decrease that reverses direction, it is perceived as a reversing trend (for up to a few reversals). If there are more than a few reversals, it appears to be cyclic or noisy data, and no trend is perceived.

Plotting multiple variables on a timeline (as in a multiple line chart) sometimes allows the user to see counter-trends. For example, if most of the variables are generally increasing and a few are decreasing, the decreasing variables can pop out and be perceived as counter-trends. If there is not much variation for any variable, it is possible to fit a regression line or curve and plot it as a *trend line* or *trend curve*. More formally, trend estimation is a statistical technique for identifying these trend lines or trend curves [5]. For purposes of discussion in this paper, we will focus only on informal trends that can be perceived visually without statistical trend estimation.

The simple approach described above only works for a number of variables along one dimension plotted against another dimension (usually time). What is the best way to see trends in two or three dimensions simultaneously?

Gapminder Trendalyzer [8] is an animated bubble chart designed to show trends over time in three dimensions. Both the size and locations of bubbles smoothly animate as time passes. This technique appears to be very effective in presentations, where a presenter tells the observer where to focus attention. It makes the data come to life, and emphasizes the critical results of an analysis. This has been done with large screens and audiences, but is probably true even for an individual presenting results to another individual; the point is that the presenter knows what is about to happen and directs the observers' attention to an area of interest. However, during analysis or data exploration, there is no presenter telling the analyst where to look. In practice, this means the analyst must replay the animation several times to identify anomalies in the trends. So, this approach may be less effective for analysis and data exploration.

This paper proposes two alternatives to animated bubble charts

---

- *George Robertson, Roland Fernandez, Danyel Fisher, Bongshin Lee are with Microsoft Research, E-Mail: {ggr, rfernand, danyelf, bongshin}@microsoft.com.*
- *John Stasko is with Georgia Institute of Technology, E-Mail: stasko@cc.gatech.edu.*

for visualizing trends in multiple dimensions, and describes a user study that evaluates the three approaches for both presentation and analysis. We are interested in understanding how effective these visualizations are for users, both as observers of a presentation and as analysts.

## 2 MULTI-DIMENSIONAL TRENDS: GAPMINDER TRENDALYZER

Gapminder Trendalyzer was created by Hans, Ola, and Anna Rosling in 2003 as a technique for using animation to illustrate trends in multi-dimensional data. Trendalyzer uses a bubble chart to show three dimensions of data, one for the X-axis, one for the Y-axis, and one for the bubble size, animated over changes in a fourth dimension (time). For example, when looking at UN statistics for various countries, the X-axis might show life expectancy, the Y-axis might show infant mortality rate, and the bubble size might show population size, with each bubble representing a country. Figure 1 shows three sample frames from an animated bubble chart similar to Gapminder Trendalyzer. The trend over time is shown as an animation over time, with the bubbles changing position and size to indicate the current data values for each country at a particular time. In the case illustrated in Figure 1 and Video Figure 1, the animation shows a general trend for most countries to increase life expectancy while decreasing infant mortality rate. However, several anomalies pop out during the animation. For example, Rwanda's life expectancy starts decreasing rapidly in 1990; this is shown in the fourth frame of Figure 1 with Rwanda highlighted.

Hans Rosling used this technology to make presentations at TED (Technology, Entertainment, Design) 2006 [16] and TED 2007 [17], evoking a strong favourable response from the audiences. This technique allows the observer to see trends in the informal sense: they can observe the general direction of movement of data over time. That is, there is no formal trend estimation. This is a very dramatic way to show trends, especially in a presentation. When Hans Rosling uses it, he is telling a story about the data and at key points in the presentation primes the observer to look at a particular part of the bubble chart before some significant event occurs. The effect adds a sense of excitement to the data: the movement of the bubbles becomes a critical part of the story.

Others have copied this approach. MicroStrategy has an Animated Bubble Chart [12] that adds the ability to collapse related bubbles into an aggregate bubble (*e.g.*, show one bubble for a continent). This aggregation technique reduces clutter and occlusion, but anomalies of interest are potentially hidden from view. Report Portal has a Moving Bubble Chart [13] which adds the ability to identify which dimensions of a data cube to map to which axes.

These techniques appear to work well for presentation of a modest number of data points (perhaps up to about 200), but several
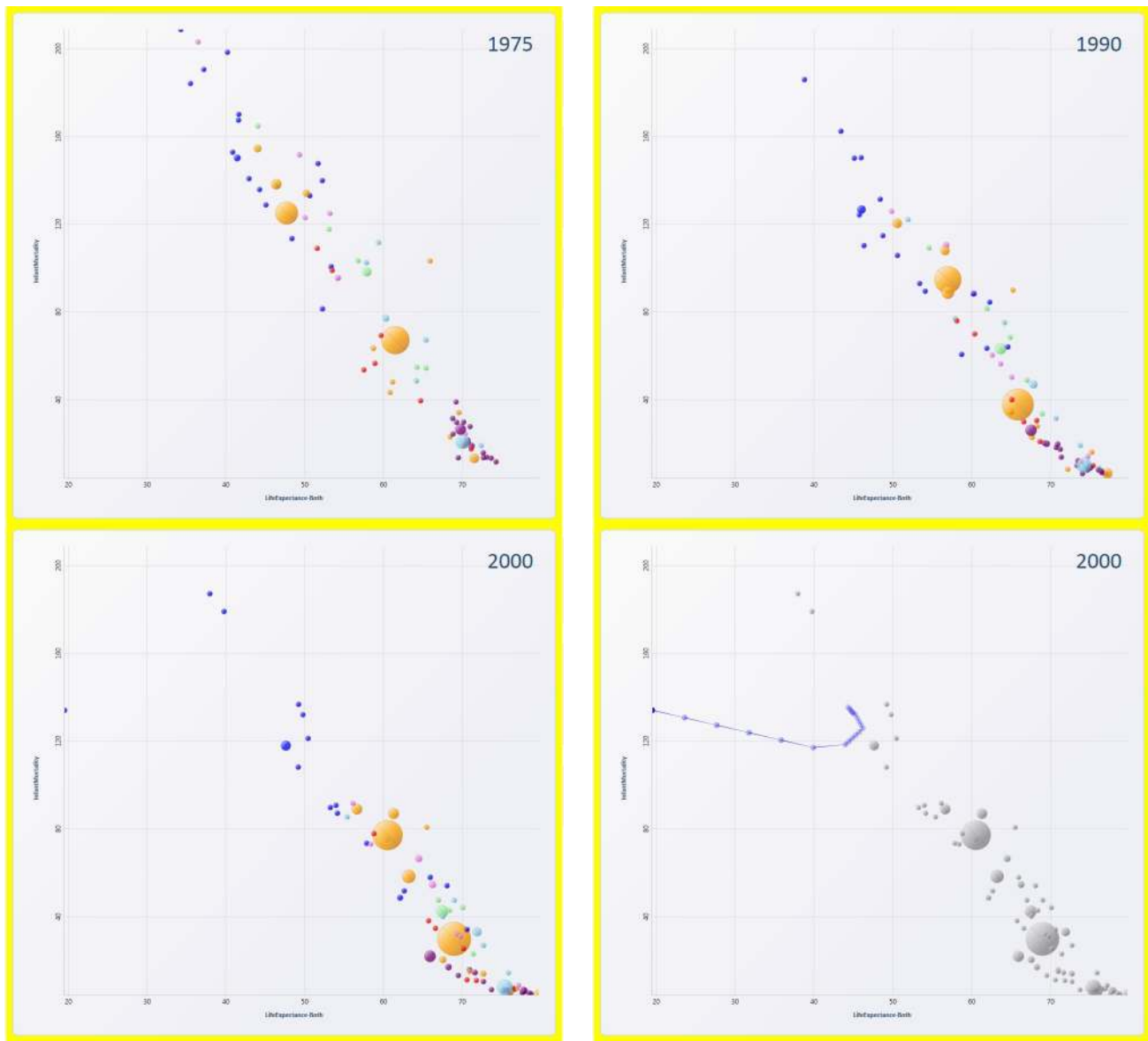
Fig. 1. Three frames of animation from the Trendalyzer-like animation tool used for the study. In the tool, the image smoothly animates between images at one year intervals. All images share the same axes. Note the convergence of most countries moving toward the bottom-right corner: high life-expectancy, low infant mortality. The fourth frame highlights one country (Rwanda) which has a counter-trend.

questions remain, which we explore in the remainder of this paper. What effect does the dataset size have? Does this technique work for analysis, where the user does not know where to look initially? Are there alternative trend visualizations that are more effective under some conditions, and what are those conditions?

## 3 ALTERNATIVE MULTI-DIMENSIONAL TREND VISUALIZATIONS

In preliminary experiments, we found that one key aspect of presentations featuring animation was having a narrator explaining what the data was doing, and highlighting important features. By directing the audience's attention, the presenter can encourage them to follow the data points of most interest. In contrast, when animation is used for analysis on unfamiliar data, the analyst does not know what points will be salient, and so will play the animation hoping that something pops out. The analyst may see several areas that look interesting, and will replay the animation several times focusing on each of the potentially interesting areas. This is a potential problem, perhaps making trend animation slower and more

error prone for analysis. If there is a lot of variability in the data, there will be a lot of random motion, making it hard to perceive trends. If there are too many data points, there will tend to be a lot of clutter and counter-trends will be harder to observe. In this section, we describe two alternative trend visualizations that may solve some of these potential problems.

### 3.1 Traces Visualization

Trendalyzer provides the user with the ability to select particular bubbles and have the animation show a trace line for each of the selected bubbles as it progresses (as shown in the fourth frame of Figure 1). This is very helpful in verifying apparent anomalies (*e.g.*, counter-trends) seen during an animation. Imagine turning all of the trace lines on at the end of the time sequence. You would see all of the trace lines simultaneously, with most anomalies popping out in that single view. This is the basis for the Traces Visualization shown in Figure 2. The Traces visualization shows bubbles at every one of their x, y locations and sizes over all time. They are then connected with edges to help clarify their sequence.
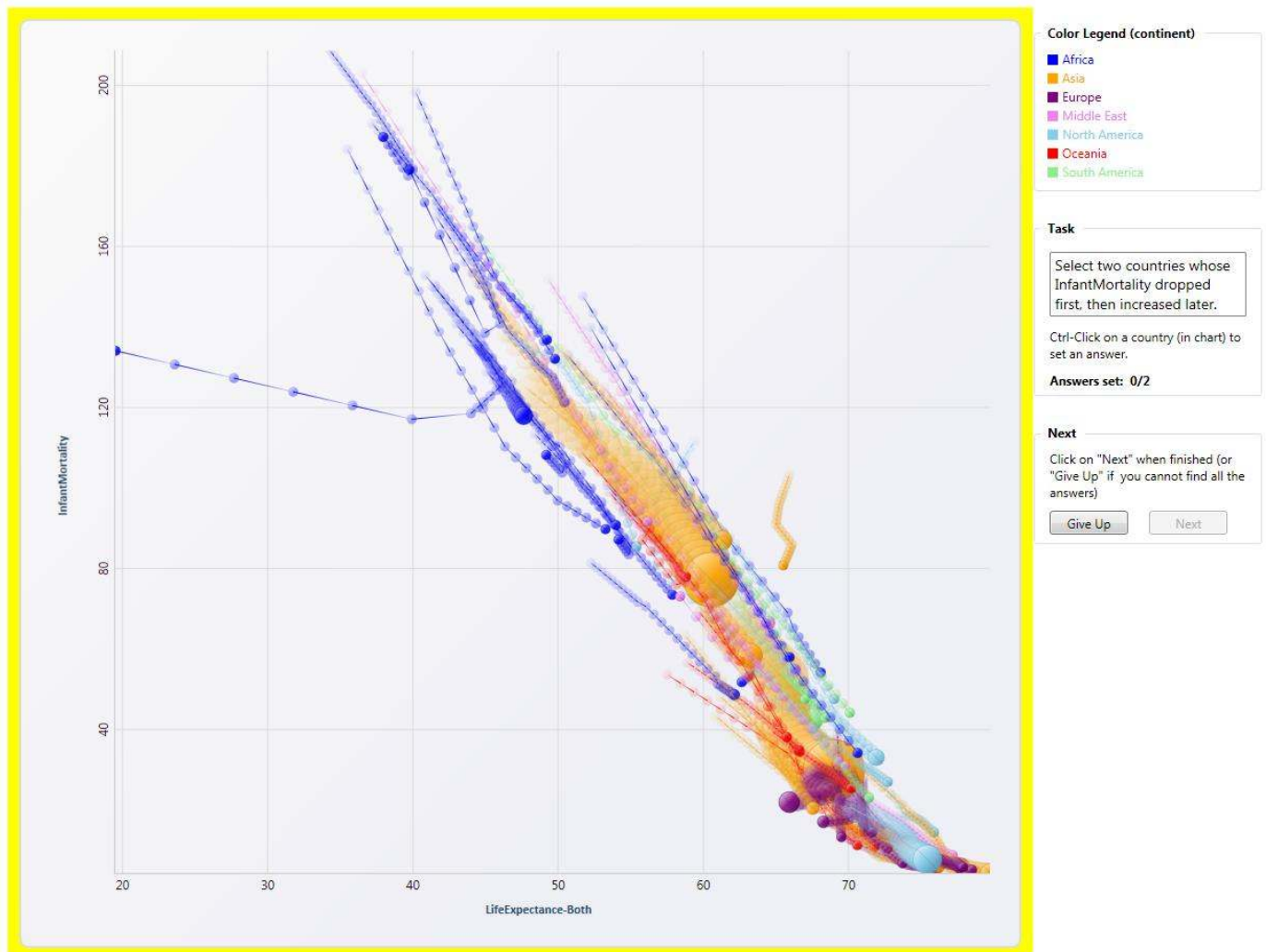
Fig. 2. Traces Visualization shows all trace lines simultaneously.

Anomalies pop-out and can be examined without animation, so analysis may perhaps be faster and less error prone. Points that move continuously through a range of values appear as clear trend lines.

One key challenge must be addressed in the design of this view. The direction of flow of the trend must be made visually apparent, since there is no animation to indicate direction. We solved this problem by using transparency of bubbles, fading from mostly transparent in the earliest to mostly opaque in the latest bubbles of the sequence. We did the same thing with the lines connecting bubbles, so that it is possible to perceive direction of flow even for smaller bubbles further apart. This can be seen in Figure 2 with Rwanda, the country on the far left, which has rapidly decreasing life expectancy. In addition, it was necessary to render larger bubbles first, to avoid occluding smaller bubbles.

One potential problem with Traces is that counter-trends that occur in the midst of many other trends may be lost in the clutter and be hard to discern. Indeed, all lines that have the same, or similar, data (even if offset in time) will be drawn over each other, as many of the countries in the middle of Figure 2 are. In addition, in a static view, reversals may be occluded: it can be hard to tell if a point is retracing its steps. This problem also occurs in animation of the same cases, but to a lesser degree.

## 3.2 Small Multiples Visualization

One way to deal with clutter that occurs for both animation and traces is to use a small multiples display [18], showing a separate trace line for each item (see Figure 3). Each of the small multiple frames shares x, y, and size axes; they are further grouped by continent and ordered alphabetically within each group.

In the design of the small multiples, we also need to find a way to allow users to know which direction the line is moving. The direction of flow is no longer discernable because each trace line is now much smaller. In addition, the original encoding of bubble size is not as meaningful because the bubbles are so small. We solved this problem by changing the encoding of bubble size to encode direction of trace flow, from smallest to largest bubble, with the largest being the original bubble size encoding at the end of the sequence.

Anomalies now pop out without any clutter, so we would expect an analyst using this method to experience comparatively few errors. However, the user has to scan the entire array of small multiples to answer certain types of questions. That serial process potentially could take longer. But compared with animation, the additional time may be offset because there is no need to replay the animation to spot anomalies.

Another issue is that the total number of data points is probably limited to around 200 before some kind of aggregation is required. However, this limit is also true of animation and traces. This limit is based on informal observation of 200 data points using all three techniques. In the case of small multiples, the size of each individual display is nearly too small to effectively see the trace line. In the case of animation and traces, the clutter becomes so extreme that it becomes hard to find the anomalies.

Fig. 3. Small Multiples Visualization shows trace lines for each country separately.

# 4 ANIMATION

Before describing the study of these three trend visualizations, we will describe some issues with animation, based on prior work. Animation is a sequence of images used to convey the illusion of movement [6]. Animation is inherently a change of some visual representation over time, hence is naturally used to facilitate perception of changes in some model over time. Animation has been used in user interfaces for a variety of purposes [1]. Animation in visualization has been used to show changes of several kinds: 1) transitions of data from one state to another [15][14][10], 2) transitions between one view and another [4], 3) illustration of how something works [21], and 4) trends [8][12][13].

One early use of animation in visualization was for algorithm animation. Kehoe *et al.* [11] describe a study that demonstrated that animation did help and noted that it improved motivation, making a difficult topic more approachable. This study suggests that using animation for trend understanding could be valuable.

Bartram [2] argues that animation evokes an emergent property of grouping when multiple, similar motions occur across a dense data display. Thus, animation allows a user to immediately recognize associated elements which may be widely dispersed, and this suggests that trends can be effectively portrayed with animation – if a group of data elements are all moving in generally the same direction, then a general trend will be perceived, and counter-trends can become apparent. Ware's [20] observation that the brain has a strong tendency to group moving objects in a hierarchical fashion

also supports this hypothesis. Ware also points out similar issues in flow visualization [20].

Unlike transition animations, which primarily help the user stay in context, trend animation conveys meaning. Timing is probably much longer because of the complexity of what is moving, and because the anomalies being sought are only seen in the movement. While a transition animation moves from a still view to a new still view, a trend animation must move continuously between states. In the studies report in this paper, we use 10 seconds for an animation depicting 25 years of data; as such, we animate at 2.5 years per second. This choice appears to be reasonable, since most participants in the study thought it was neither too fast nor too slow. A similar choice is made in Rosling's demonstrations. But, we do not know the optimal time for a trend animation. In our study, this is partially solved by interaction. That is, users are provided a control that allows them to pause and advance the animation as they wish.

Cleveland & McGill [7] describe how easily different graphical features are perceived and judged. Their results suggest that position is judged far more accurately than area. In animated bubble charts, this implies that changes of position of bubbles will be much more accurately judged that changes of bubble size. The implication is that it will not be easy to accurately observe trends or anomalies in bubble sizes during animation. Because of this, we did not rely on or test bubble size animation in the study reported here.

Tversky *et al.* [19] are skeptical of many published animation studies, citing problems with study methodology. In their own studies of use of animation to illustrate or communicate the workings

of complex systems, they found no benefit to animation. However, they acknowledge that the use of animation for transitions may well lead to a benefit. Their studies do cast some doubt on the effectiveness of animation for trends.

Zongker and Salesin [21] counter Tversky's argument by describing a number of design principles that make animation presentations effective.

Baudisch *et al.* [3] describe a static alternative to animated transitions with their "Phosphor" system. In one of their studies, they demonstrated a case where use of a static depiction of motion was faster than use of an animated alternative. Their strobe approach is similar to what we propose in our Traces trend visualization.

Griffin *et al.* [9] compare animated maps to static small multiples displays, finding significant advantages in both speed and accuracy for animated maps. Their small multiples display shows temporal slices, essentially showing stages of the animation as in Figure 1. In contrast, our small multiples solution is showing slices through the data, showing only one trend line in each small multiples frame.

## 5 EXPERIMENTATION

We performed a study to test the effectiveness of the three trend visualizations. We used a 3 (Visualization: Animation vs. Traces vs. Small Multiples) x 2 (Dataset size: Small vs. Large) x 2 (Use: Presentation vs. Analysis) mixed study design. The Presentation vs. Analysis variable was between subjects, while each of the other variables was a 3 (Visualization) x 2 (Dataset size) within-subjects design. The same datasets and tasks were used for both Presentation and Analysis studies. As will be described below, the interaction techniques were limited in the Presentation study.

### 5.1 Hypotheses

We postulated four main hypotheses for the experiment.

H1. Animation will be more effective than other techniques when used for Presentation and less effective than other techniques when used for Analysis. That is, participants will be (a) faster and (b) make fewer errors in the Presentation condition.

H2. Traces will be more effective than Animation when used for Analysis. That is, participants will be (a) faster and (b) make fewer errors in the Traces condition.

H3. Small Multiples will be more effective than Animation when used for Analysis. That is, participants will be (a) faster and (b) make fewer errors in the Small Multiples condition.

H4. Participants will be more effective with small datasets than with large datasets. That is, participants will be (a) faster and (b) make fewer errors when working with Small datasets.

### 5.2 Interaction Techniques

To compare the performance of our three alternatives for trend visualization, it is important to describe the interaction techniques supported for each, and to ensure that differences in interaction techniques are understood with regards to the outcome of the study.

For Animation, there is a Play button in the lower left corner below the bubble chart (see Video Figure 1). When that button is pressed, a 10 second animation of the changes over time is played. During play, the button changes to a Pause button. To the right of the button is a time slider. The user can grab the time slider control and move to any point in time, moving the slider at any speed to control playback speed (instead of the default 10 second playback). One or more bubbles may be selected by clicking on them, and deselected by clicking a second time. When one or more bubbles are selected, selected bubbles show a trace, while unselected bubbles are greyed out and show no trace. This makes the selected bubbles more salient, and deals with clutter issues.

For Traces, there is no Play button or time slider. Selection works as in Animation, but unselected traces are greyed and made more transparent. This allows the selected traces to be rendered more salient, and help deal with clutter issues.

### 5.3 Datasets

All data came from the United Nations Common Database (http://unstats.un.org/unsd/cdb/) which provides historical data by country. We downloaded various tables for approximately 270 countries for data from 1975 through 2000. The tasks, which will be described in the next section, were inspired by questions addressed by Hans Rosling in his TED conference presentations [16][17]. In order to control the time for each task, we pared down the data to two study datasets, one with 80 countries and one with 18. These were chosen so that overall trends in the smaller datasets matched overall trends in the full dataset.

The datasets had missing data (*i.e.*, some year's values were missing from the UN data). In cases where the missing data was at the beginning or end, data was extrapolated from nearby data. In other cases, data was filled in with interpolated data. For the purposes of the study, it was not important that the data be entirely accurate. Participants were told to answer questions based on what they saw, not on any preconceived notions or beliefs about countries.

### 5.4 Tasks

The experiment included 24 tasks (see Table 1), each with one to three required answers. In some cases, there were more than three correct answers possible. Nine tasks had continents for answers and 15 had countries for answers. There were 8 continents, so continent questions required choosing 1-3 out of 8. Country questions required choosing 1-3 out of either 18 or 80. In either case, random guessing was unlikely to provide a correct answer. Participants were asked to work as quickly and accurately as possible.

Table 1. Sample of tasks used in study.

| |
|---|
| Select 3 countries whose rate of energy consumption was faster than their rate of GDP per capita growth. |
| Select 2 countries with significant decreases in energy consumption. |
| Which country had the most significant decrease in GDP per capita? |
| Which continent had the most significant increase in GDP per capita (i.e., the continent with largest percentage of countries with significant increases in GDP)? |
| Select 2 countries whose GDP per capita increased first, then decreased later. |
| Which continent had the least changes in GDP per capita? |

Bubbles were color-coded by continent, and a color legend appeared in the upper right (see Figures 2 and 3). If bubbles were selected, their colors remained and all other bubbles changed to grey.

Participants selected an answer by control-clicking on either a country bubble or a continent (in the color legend). There was a yellow highlight around the area where the answer was expected (*i.e.,* either around the bubble chart or around the color legend). There was a "Give Up" option, but it was only used 2 times out of 864 (24 x 2 x 18) trials. There was no time limit.

Task and visualization order were both counterbalanced. The order of dataset sizes in each study was fixed, with the smaller dataset used first for each visualization. This allowed participants to build their skills as they proceeded with the tasks over time.

### 5.5 Study Method

Both Analysis and Presentation conditions used a 3 (Visualization) x 2 (Dataset size) within-subjects design. Tutorials and tasks were the same for both studies, and were presented automatically. Time and error measurements were captured automatically. For each visualization, tutorial tasks were provided (three for the first and two for the second and third), followed by eight tasks (four small dataset tasks followed by four large dataset tasks), followed by survey questions specific to that visualization. A set of general survey questions was provided at the end of the session. Each session lasted

about 1.5 hours. A software gratuity was provided for participation. Participants were screened to ensure that they were not color-blind, had general computer experience, understood scatter plots, and were balanced for age and gender.

### 5.5.1    Analysis Condition

The intent of the Analysis condition was to examine how people interacted actively with the visualization, simulating what an analyst does to discover insights in trends. Each participant used a 2.4 GHz Dual Core Pentium PC with 3.5 Gb of RAM running Windows Vista. Each PC had a 21" monitor running at a resolution of 1280 x 1024. Two participants were run at a time, although they did not interact with each other.

For each visualization, users practiced with a set of simple tutorial tasks intended to help them understand the visualization and the interaction methods. The participants had full use of the interaction techniques described earlier; in particular, in Animations they were able to play the animation as often as they wished. Users also had the ability to select traces in order to highlight them.

There were 18 subjects (8 male, 10 female) with an average age of 38 (minimum 18, maximum 52).

### 5.5.2    Presentation Condition

The intent of the Presentation condition was to examine how people react when *passively* exposed to the visualization, simulating a conference or presentation. We could not fully simulate the experience of hearing a talk of course; as a real speaker would actually describe the fact they wanted users to note. Our approximation was to have the speaker highlight a relevant fact that was meant to guide the listener toward the right answers. The listeners would then answer the actual task without guidance.

Six participants at a time viewed a scripted presentation of tasks on a projected display (approximately six feet diagonally). For non-animation tasks, the presentation was a recording made with Camtasia on a 2.4 GHz Dual Core Pentium PC with 3.5 Gb of RAM running Windows Vista. The animation tasks were not pre-recorded, but were presented on a similar computer by a presenter following a script. In all three cases, all (three groups of six) participants heard the same script and saw the same things on the display.

For all tasks and views, the presenter described a trend that was relevant to the task, but different. Table 2 has an example of this relevant trend description for Task 15, "Which continent had the least changes in Infant Mortality?" In each view, the presenter would point to the countries or trends being described. For the animation tasks in particular, the animation was shown once during the presenter's narration, with those countries selected. The relevant countries were highlighted during the animation. The selected countries were then deselected and the animation was reset to the first date in the sequence. The task was then described and the animation was played a second time with nothing selected.

Table 2. The narration for task 15 in the presentation condition.

> **Task 15:** "On the X axis, Life Expectancy--a higher number means that people can expect to live longer. On the Y axis, Infant mortality--which means lower down is better. As you can see, most countries seem to move from the top-left (high infant mortality, low life expectance) toward the bottom right (low infant mortality, and living longer)--but at different speeds. In 2000, India's life expectancy and infant mortality have caught up to where China was in 1975."

Each participant in the Presentation condition had a Pentium laptop running the same software as in the Analysis condition. However, in Presentation condition, interaction techniques were disabled. For the Animation view, the participants only saw the closing state of the animation, exactly as it appeared on the presentation screen at the end of the presentation. Participants were only able to enter answers to the questions by control-clicking on either a country bubble or continent in the continent list (which would highlight the selected answer). They were not able to select

countries; they could not control animation playback. The participants did not interact with each other or the presenter. After the final survey, an informal group discussion focussed on what worked well and what did not.

There were 18 subjects (10 male, 8 female) with an average age of 40 (minimum 23, maximum 53).

## 5.6    Study Results

We present results from the study in three parts; accuracy, completion time, and subjective preferences. A 3 (Visualization: Animation vs. Traces vs. Small Multiples) x 2 (Dataset size: Small vs. Large) x 2 (Use: Presentation vs. Analysis) Repeated Measure Analysis of ANOVA (RM-ANOVA) was performed for each metric. Only significant results are reported. We corrected for multiple tests in posthoc analyses using the Bonferroni technique.

### 5.6.1    Accuracy

Since some of the tasks required multiple answers, accuracy was computed as a percentage. For example, if the participant provided only one right out of two correct answers, we counted it as 50% accurate rather than a wrong answer. Overall, accuracy was quite low in this study, averaging 65%, indicating that the tasks were more difficult than we anticipated.

There were two significant accuracy results. First, we observed a significant main effect of Visualization ($F_{2,68}=4.18$, p=.029), with posthoc tests showing that Small Multiples was significantly more accurate than Animation (p<.001) The Traces condition was not distinguishable from Animation or Small Multiples conditions (Figure 4).
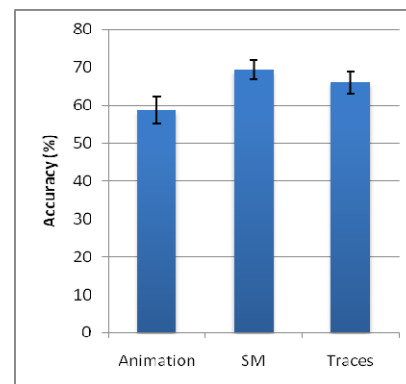


Fig. 4. Small Multiples was more accurate than Animation, p<.001.

Second, participants were more accurate with the Small dataset ($F_{1,34}=21.40$, p<.001, see Figure 5). This supports hypothesis H4.b.
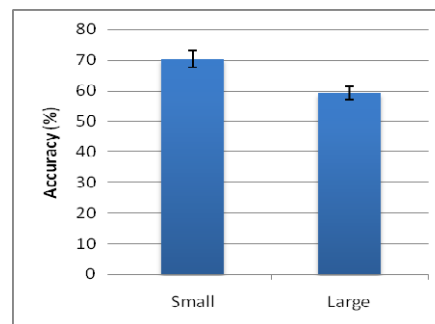


Fig. 5. Small was more accurate than Large, $F_{1,34}=21.40$, p<.001.

### 5.6.2    Task Completion Time

An answer was deemed incorrect if the participant provided none of the correct answers (i.e., 0% accuracy). We analyzed the data both

with and without errors, and the pattern of significant results the same. We will show the results excluding the errors. These results were analyzed in log time to control for the skewness in reaction time data, but are reported in seconds for clarity.

In terms of time to task completion, we observed a significant interaction effect between Visualization and Use ($F_{2,64}=42.82$, $p<.001$), with posthoc tests showing that Animation was fastest in Presentation and slowest in Analysis. First, in Presentation mode, Animation was *faster* than Small Multiples ($p<.001$) and Traces ($p<.001$). The mean time for Animation was 15.80 seconds compared to 25.30 seconds for Small Multiples (about 60% faster), and 27.80 seconds for Traces (about 76% faster). Traces and Small Multiples were not statistically distinguishable.

Second, in Analysis, Animation was *slower* than Small Multiples ($p<.001$) and Traces ($p<.001$) The mean time for Animation was 83.10 seconds compared to 45.69 seconds for Small Multiples (about 82% slower) and 55.01 seconds for Traces (about 51% slower). This supports hypothesis H2.a and H3.a, Traces and Small Multiples are both faster than Animation when used for Analysis.

These two observations support hypothesis H1.a, Animation is faster when used for Presentation and slower for Analysis.

### 5.6.3    Subjective Preferences

At the end of each set of trials for a particular visualization, there was a survey with the following seven questions. Each were answered with a 7-point Likert scale (0 = strongly disagree, 6 = strongly agree). We ran a 3 (Visualization) x 2 (Use) Multivariate ANOVA and found a significant effect of Visualization. See Table 3 for pair-wise comparisons.

Table 3. Average ratings for seven questions for each visualization. * indicates significant differences ($p<.05$).

|  | Animation | SM | Traces |
|---|---|---|---|
| Q1. The visualization was helpful to me in answering the questions. | 4.6 *Traces | 4.2 | 4.1 |
| Q2. For the smaller dataset, I found the tasks easy using this visualization. | 4.6 *SM | 4.2 | 4.5 |
| Q3. For the larger dataset, I found the tasks easy using this visualization. | 2.6 | 3.4 *Traces | 2.3 |
| Q4. I enjoyed using this visualization. | 4.3 *SM *Traces | 3.7 | 3.5 |
| Q5. I found this visualization exciting. | 4.3 *SM *Traces | 3.1 | 3.0 |
| Q6. For the smaller dataset, I found the screen too cluttered. | 1.8 | 1.5 | 2.0 |
| Q7. For the larger dataset, I found the screen too cluttered. | 4.4 | 2.8 *Animation *Traces | 4.7 |

The significant differences indicate that Animation was judged to be more helpful than Traces. Animation was judged to be easier than Small Multiples for small datasets. Small Multiples was judged to be easier than Traces for large datasets.

Animation was judged to be more enjoyable and more exciting than both Traces and Small Multiples.

Small Multiples was judged to be less cluttered for large datasets than both Animation and Traces.

At the end of all trials, there was a general survey with 23 questions. Again, each was answered with a 7-point Likert scale (0 = strongly disagree, 6= strongly agree). Table 4 shows a few of the general questions, along with their average ratings.

Low average ratings for G7 and G8 indicate that participants found the animation speeds neither too fast nor too slow.

The high rating on G9 indicates that users tended to lose track of data points as they moved.

Table 4. Average ratings for a few general questions.

|  | Presentation | Analysis | Overall |
|---|---|---|---|
| G1. I found the Traces view enjoyable. | 3.8 | 2.9 | 3.4 |
| G3. I found the Small Multiples view enjoyable. | 4.1 | 3.4 | 3.7 |
| G5. I found the Animation view enjoyable. | 4.6 | 5.0 | 4.8 |
| G7. The animation went too fast for me. | 3.2 | 2.8 | 3.0 |
| G8. The animation went too slow for me. | 1.6 | 1.3 | 1.4 |
| G9. I lost track of some data points as they moved. | 4.9 | 4.6 | 4.8 |

Finally, general survey questions G13 to G23 are subjective preference questions. The most interesting result is from G13 and G14; general preferences are shown here with their raw counts.

G13: Which visualization did you PREFER for the small dataset?
G14: For the large?

Presentation, small: Animation (9) > SM (6) > Traces (3)
Presentation, large: Traces (8)   > SM (6) > Animation (4)
Analysis, small:   Animation (7) > SM (6) > Traces (5)
Analysis, large:   Animation (8) > SM (6) > Traces (4)

Animation received slightly more votes, except for presentation with Large datasets, where Traces received slightly more votes. This is particularly interesting, because this is where Gapminder appeared to excel.

After the presentation, the groups of users were invited to discuss their impressions of the visualizations. Interestingly, while each group came to consensus, they disagreed on which was the best choice: one group preferred Animation, one preferred Traces, and one preferred Small Multiples.

In these final discussions, participants often commented on the difficulty of tracking objects in animation, particularly for large datasets. With no control over the animation, they had to find their target very rapidly: "It's hard to check all the flying balls!" However, participants in all groups did find flying balls to be exciting.

One user complained about reversals: "Sometimes, a country is going up, and then it goes down, and I realize I've been looking at the wrong thing."

On the use of Traces on large datasets, one user reported: "That's confusing—and blurry. After staring at it for a few minutes, I can now see some trends that I couldn't before."

Even for static displays, participants preferred to look for data on the large shared screen rather than the small screen on their laptops. This reassures us that users were, in fact, thinking about the problem in terms of a presentation, rather than studying their own display.

### 5.7    Discussion

In trying to understand the results, there are a number of issues that must be discussed. In the Presentation condition, Animation was fast because the users could not replay the animation; as a result, they were forced to already have made a decision (or to guess) during the first run. The only data on the screen was the final year, and so they either had a correct solution or gave up. For the other views, by design more information was available on the screen. In contrast, Animation in Analysis mode gave users extensive interactive tools,

and many users spent a long time exploring this data, replaying the animation. Despite this extra time, the accuracy rate for analysis of animation was not significantly higher than presentation of animation. Thus, hypothesis H1.a was supported, but H.1.b was not.

For both Presentation and Analysis, users rated the animation as *fun*, but small multiples appeared to be *more effective*.

Hypothesis H2.a was supported; Traces is faster than Animation in Analysis. However H2.b was not supported; there was no significant difference in error rates.

Hypothesis H3.a and H3.b were supported; Small Multiples is faster and leads to fewer errors than Animation in Analysis.

One completion time result was significant only when considering all data, including errors. Participants were faster with the Small dataset ($F_{1,34}$=5.49, p=.025). The mean time for Large was 45.96 seconds and for Small was 43.20 seconds, or about 6% faster. Post-hoc analyses using Bonferroni corrections showed that Small Multiples accounted for the result ($t(35)$=-6.58, p<.001). This weakly supports hypothesis H4.a.

Hypothesis H4.b was supported; participants had fewer errors with the small dataset than with the large dataset.

All three techniques fail to scale beyond about 200 data points. One way to handle larger scale is to introduce abstraction, as MicroStrategy's Animated Bubble Chart [12] does. However, this is likely to hide anomalies and make analysis potentially harder.

Accuracy is an issue for trend visualization, especially for Large datasets. Animation and Traces tend to become cluttered as the size of the dataset increases. Small Multiples removes clutter, but each small multiple frame becomes smaller as dataset size increases.

During our after-study discussions, users repeatedly reported that the animation condition was "fun" or "exciting" or, in one case, "emotionally touching." Yet they also found it confusing: one user complained that "the dots flew everywhere," and shook her head in frustration. One user pointed out a case when a country's trend reversed itself partway through, and complained: "I was following it until then, but then I lost it."

## 6 CONCLUSION

Gapminder Trendalyzer [8] introduced an exciting way to use animation in presentations to show trends in data in multiple dimensions. This paper has compared trend animation with two alternative trend visualizations, both using static depictions of trends.

Earlier work by Tversky *et al.* [19] and Baudisch *et al.* [3] hinted that static depictions of motion can be as effective as or more effective than animation in some cases. The study reported in this paper partially supports that.

Trend animation may be successful in presentations in practice. However, in our studies, users appeared to find the movement confusing, even when guided. Presenters showing animated data should strongly consider ensuring that their data tells a clean story: this can be confused by having too many data points, by data points that reverse their tracks over time, or by having points that do not move in synchrony.

Traces and Small Multiples worked best for analysis. The study supports the intuition that using animation in analysis requires replaying the animation multiple times to discover where to focus. The study also suggests that Small Multiples leads to fewer errors.

Despite the subjects finding trend animation to be more fun and exciting, they only slightly preferred it to the static depictions. Using animation to show trends in data clearly has value, especially for presentation. However, for analysis, static depictions of trends appear to be more effective.

What does this study imply for an analyst or a presenter? As we begin to look at time-varying, high-dimensional data, we need to rethink the relationship between analysis and presentation. Historically, a scatter plot is as valuable for analysis as for presentation afterward. This study suggests that this may no longer be true. An analyst should consider examining their data with a small multiples view, but then presenting it with animation. A presenter must ensure that observers do not feel overwhelmed by the data. In part, this depends on the behaviour of the data: points that follow trends smoothly will be easier to follow then those that reverse themselves or jump around. To the extent that the data fits the presenter's story well, the visualization will be easier to understand.

## REFERENCES

[1] R. Baecker and I. Small, Animation at the Interface, *The Art of Human-Computer Interface Design*, B. Laurel, editor, Addison-Wesley, 1990, pp. 251-267.

[2] L. Bartram. Enhancing Visualizations with Motion. In *Hot Topics: Information Visualization 1998*, May 1998.

[3] P. Baudisch, D. Tan, M. Collomb, D. Robbins, K. Hinckley, M. Agrawala, S. Zhao, and G. Ramos. Phosphor: Explaining Transitions in the User Interface Using Afterglow Effects. In *Proc. ACM UIST 2006*, Montreux, Switzerland, Oct 2006.

[4] B. Bederson and A. Boltman. Does Animation Help Users Build Mental Maps of Spatial Information? In *Proc. IEEE InfoVis 1999*, San Francisco, CA, Oct 1999.

[5] M. Bianchi, M. Boyle, and D. Hollingsworth, "A comparison of methods for trend estimation", *Applied Economics Letters*, 6(2): 103-109, 1999.

[6] Computer Animation, http://en.wikipedia.org/wiki/Computer_animation.

[7] W. Cleveland and R. McGill. Graphical Perception and Graphical Methods for Analyzing Scientific Data. *Science*, 229:828-833. 1985.

[8] Gapminder. http://www.gapminder.org

[9] A. Griffin, A. MacEachren, F. Hardisty, E. Steiner, and B. Li, A Comparison of Animated Maps with Static Small-Multiple Maps for Visually Identifying Space-Time Clusters, *Annals of the Association of American Geographers*, 96(4), 2006, pp. 740-753.

[10] J. Heer and G. Robertson. Animated Transitions in Statistical Data Graphics. IEEE Transactions on Visualization and Computer Graphics, vol. 13, no. 6, pp. 1240-1247, Nov/Dec, 2007.

[11] C. Kehoe, J. Stasko, and A. Taylor, Rethinking the Evaluation of Algorithm Animations as Learning Aids: An Observational Study, *International Journal of Human-Computer Studies*, 54, 2, 2001, pp. 265-284.

[12] MicroStrategy Animated Bubble Chart, http://www.microstrategy8.com/DashboardGallery/Dashboards/AnimatedBubble.html.

[13] Report Portal Moving Bubble Chart, http://www.reportportal.com/Flash/Bubble/Bubble.html.

[14] G. Robertson, K. Cameron, M. Czerwinski, and D. Robbins. Animated Visualization of Multiple Intersecting Hierarchies. *Journal of Information Visualization*, **1**(1):50-65. Palgrave, 2002.

[15] G. Robertson, S.K. Card, and J.D. Mackinlay. Cone Trees: Animated 3D Visualizations of Hierarchical Information. In *Proc. ACM CHI 1991*, New Orleans, LA, Apr 1991.

[16] H. Rosling, TED 2006, http://www.gapminder.org/video/talks/ted-2006---debunking-myth-about-the-third-world.html.

[17] H. Rosling, TED 2007, http://www.gapminder.org/video/talks/ted-2007---the-seemingly-impossible-is-possible.html.

[18] E. Tufte. *Envisioning Information.* Graphics Press, 1990.

[19] B. Tversky, J. Morrison, and M. Betrancourt. Animation: Can It Facilitate? *Int. J. Human-Computer Studies*, 57:247-262, 2002.

[20] C. Ware, *Information Visualization, Perception for Design*, Morgan Kaufmann, 2004.

[21] D. Zongker and D. Salesin. On Creating Animated Presentations. In *Proc. Eurographics/SIGGRAPH Symposium on Computer Animation.* 2003.