Audio Enhanced Map Animation – Testing Perception of Spatial and Temporal Relations

Andreas Hall, Salla Multimäki, Paula Ahonen-Rainio

Aalto University, School of Engineering, Espoo, Finland

Abstract. Map animation is a popular way to represent time in geovisualization but studies of the effectiveness of map animations have produced mixed results. One potential way to improve map animations that has been discussed frequently but only sparsely studied in the literature is to add sound and make them multimodal. Therefore, this paper investigates whether sound can be used to enhance the usability of an animated map in an exploratory analysis context. Here, we focus on the perception of spatial and temporal relations between events. An internet-based user test was performed. The results show that adding sound does not impair the usability of map animations but fail to show that it improves it. The conclusion is that, rather than rejecting the hypothesis that sound can be used to improve map animations, the study was not appropriately designed to measure what it set out to measure. However, it produced some interesting results indicating that the benefit of sound in animated maps might be related to the number of events in the animation and be different when it comes to reasoning about temporal and spatial relations. This study works as a good starting point for future research.

Keywords: Map Animation, Sonification, Audio Temporal Legend, User Testing, Multimodality

1. Introduction

The field of geovisualization deals with both spatial and temporal data. Temporal data, i.e., data that embody change over time, are challenging for all disciplines related to data visualization and analysis (Andrienko et al. 2010, Kriglstein et al. 2014). In recent years, map animation has become increasingly popular as a way to represent time in geovisualization (Harrower & Fabrikant 2008) and it is included in almost all the geovisualization software packages that deal with spatio-temporal data.

Map animation can use time itself to represent time, unlike static maps, which have to use space to represent time. Studies of the effectiveness of animated versus static maps have produced mixed results (e.g., Slocum et al. 2001, Tversky et al. 2002, Griffin et al. 2006, Harrower & Fabrikant 2008, Aigner et al. 2011, Kriglstein et al. 2014). Among the benefits of map animation are its ability to congruently represent the passage of time and the fact that the representation of time through time itself makes more screen space available to represent non-temporal data. Among the drawbacks are that users cannot see a phenomenon at a glance any more, that the cognitive load on short-term memory is greatly increased, and that change blindness (Simons & Rensink 2005) becomes a factor. Further research is still needed for animated maps to reach their full potential, e.g., about how variations in the design of animated maps affect our ability to communicate and learn (Harrower & Fabrikant 2008).

Although geovisualization primarily comprised only visual means, more than a decade ago several authors (Fairbairn et al. 2001, MacEachren & Kraak 2001, Slocum et al. 2001) pointed to the need for multimodal approaches in interface design. In psychology, it is well known that our senses work together to produce a representation of the external world that “enables us to function better by exploiting the redundancies and complementarities provided by multiple sensory modalities” (Alais et al. 2010, p. 3). Therefore, multimodal map animations, if properly designed, can be expected to function better than unimodal map animations.

Sound is especially well suited to making map animations multimodal as sound is an inherently temporal phenomenon. The use of sound in animations is also motivated by findings in cognitive science and related disciplines. Sound can reduce split attention as it enables us to benefit from our innate dual-channel capabilities to both hear and see at the same time, it can reduce cognitive overload as it can be used to reduce the amount of information that needs to be processed visually, and it can be used to make users more cognitively active by having them make connections between visual and auditory stimuli (Mayer 2005, Harrower 2007).

Using sound to convey information is known as *sonification*. It can be seen as the auditory counterpart of data visualization (Reuter et al. 1990), addressing the “transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation” (Kramer et al. 1999, p. 3). Dubus and Bresin (2013) give a systematic review of work related to the sonification of physical quantities.

Sonification has been present in the field of geovisualization since Krygier (1994) proposed a set of abstract sound variables for the representation of spatial data analogous to Bertin's (1983) graphic variables. It has been studied in relation to a wide range of applications, e.g., virtual environments (e.g. Fröhlich et al. 1999), mapping for the visually impaired (e.g. Zhao et al. 2008), and systems for the exploration of geospatial data (e.g. Harding et al. 2002). Recent examples of sonifications include the use of sound to enhance the map use experience (Laakso & Sarjakoski 2010) and to represent uncertainty in data (Bearman & Lovett 2010), a topic discussed already by Fisher (1994). However, very little research has been done on using sound in map animations. Krygier (1994) reported two applications, one where loudness was used to add additional information to an animated map and one where pitch was used to replace the time bar, i.e., the temporal legend, in an animated map. More recently, several authors have suggested that map animations could benefit from audio-temporal legends (Kraak et al. 1997, Slocum et al. 2001, Midtbø 2001, Harrower 2007, Harrower & Fabrikant 2008). This also seems to be supported by findings in neuroscience indicating that hearing shapes our perception of time (e.g. Heming & Brown 2005, Kowalska & Szelag 2006, Bolognini et al. 2012). However, no systematic testing of the use of audio-temporal legends can be found in the literature.

Therefore, in this study we investigated whether sound enhances the usability of an animated map in an exploratory analysis context. The research questions were the following. Does an audio-enhanced animated map facilitate the analytical reasoning of an analyst? Furthermore, does audio enhancement have a different effect on the perception of spatial and temporal relations? How is the effect of audio enhancement related to the amount of information in an animation? The hypothesis was that correctly added sound facilitates analytical reasoning by map animation, especially regarding temporal relations.

The research was done as an internet-based user test. The stimuli used were map animations that applied sound to the introduction of new point-type events, to encode thematic information about the events, and as a temporal legend. The test had a primary quantitative part evaluating *user performance* and a secondary qualitative part evaluating *user experience*, according to the scenario-based approach described by Lam et al. (2012). These were both *scenarios for understanding visualizations*.

The user test is introduced in the next section and then the results in the following section. After this, the results are discussed and ideas for future research are presented. The paper ends with a concluding section.

2. Test Settings

This section first reports the study design and the stimuli used, then the tasks and the apparatus used. After this, the test procedure is explained and there are some details about the participants.

2.1. Study Design and Stimuli

To be able to answer the research questions the study design had to differentiate not only between animations with and without sound but also between the perception of spatial and temporal relations and animations with high and low amounts of information contents. Therefore, the study applied a three-variable mixed design with *audio* (sound vs. no sound) as a between-group variable and *task* (spatial vs. temporal) and *event frequency* (low vs. high, i.e., the number of events in the animation) as the within-group variables (*Figure 1*). The dependent variable was the test participants' task performance (accuracy).

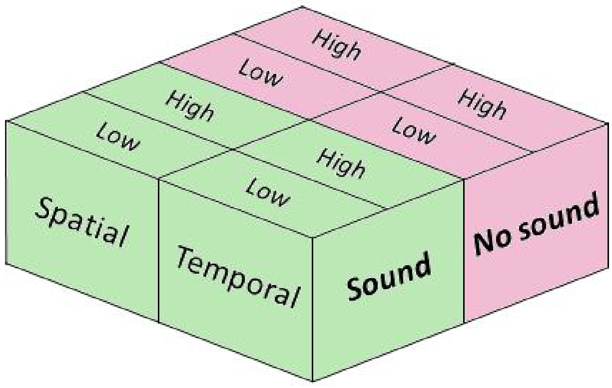

Figure 1. The study design visualized as a cube.

Four animated maps, with durations of about 20 seconds each, were used as stimuli. They were clips from a 14-minute-and-24-seconds-long multimedia artwork named "1945-1998" made by Isao Hashimoto in 2003[1] (here used courtesy of Isao Hashimoto). The artwork showed the time and location of every known nuclear detonation during the years 1945-1998 and the country that performed each of them. It was chosen as it caught the researchers' attention as an interesting case of using sound together with visual means to

[1] Available at http://www.ctbto.org/specials/1945-1998-by-isao-hashimoto/

communicate spatio-temporal information and as it had been used successfully in previous research (Hall & Ahonen-Rainio 2014). Furthermore, it presented real data[2] and was ready to use as it was.

In the animations, as in the original artwork, the locations of the detonations were plotted on a world map as points colored according to the country that performed them. Real-world time was linearly scaled to animation time and a textual temporal legend showed the current month and year as numbers. One second in the animations equaled one month in real-world time. As time went by, the detonations appeared and disappeared on the map. The total and country-wise cumulative numbers of detonations were also given at every moment in time. In the test an additional legend was shown alongside the animations that named the colors and the countries they represented as the detonations were referred to by color in the test. *Figure 2* shows a screenshot of how the animations and the additional legend was presented to the participants in the test.

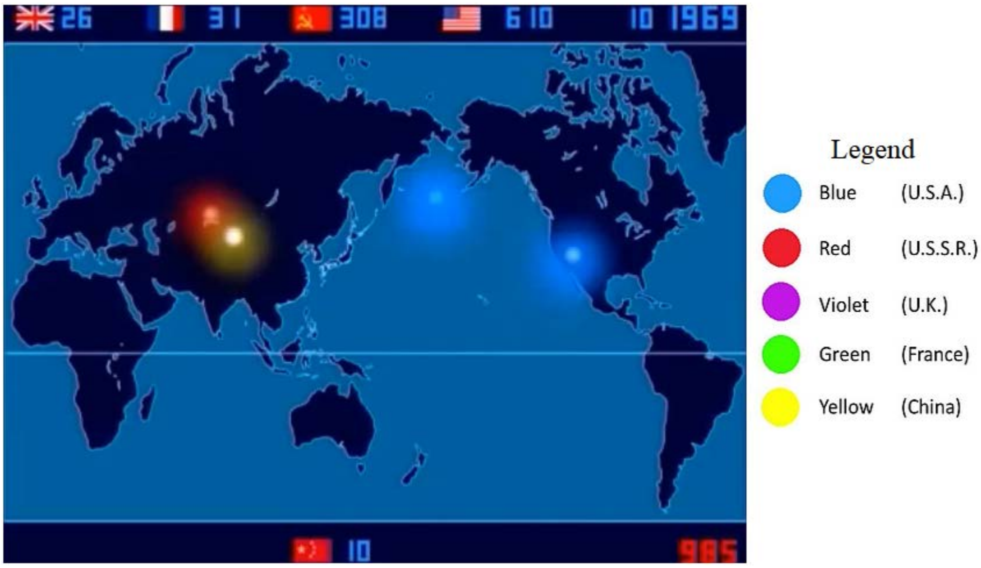

Figure 2. A screenshot of how the animations and the additional legend was presented to the participants in the test. In the upper right-hand corner of the animation is the temporal legend and in the lower right-hand corner the cumulative number of detonations performed can be found; the numbers next to the flags represent the cumulative numbers of detonations sorted by country. The additional legend is on to the right of the animation.

---

[2] Available at http://inis.iaea.org/search/search.aspx?orig_q=RN:31060372

The animations, as the original artwork, applied audio information in several ways. They had an audio temporal legend applying two kinds of metronomic beeps, one each second signaling months passing by and another one every twelfth second signaling a change of year. In addition to this, every detonation was accompanied by a beep. The beeps also included thematic information; the pitch of the beep was different for each country. However, the differences were relatively subtle.

The sound group saw the animated maps with sound and the no-sound group saw them without sound. The length of the animated maps and how they relate to the within-group variables can be found in *Table 1.* The numbers in the first column indicate the order in which the participants saw them. The table also states the number of events (detonations) occurring and the number of countries performing detonations in each animation. By alternating animations with spatial and temporal tasks the learning effect was minimized and by going from a few to many events the difficulty increased as the test went on. The start and stop times in the table refer to the time in the original artwork. The size of the animations used in the test was 640 x 480 pixels.

Animation	Task	Event freq.	Events	Start	Stop	Duration	Countries
1	spatial	low	35	2:57	3:18	21 s	3
3	temporal	low	46	3:18	3:39	21 s	3
2	spatial	high	278	4:34	4:56	22 s	4
4	temporal	high	122	6:00	6:20	20 s	5

Table 1. The four animated maps used in the test.

2.2. Tasks

The participants were asked to thoroughly watch each animated map twice and, depending on the animation, to focus on either where or when events took place (different focus on different animations) and on the color of the events. In addition, the participant was told: “Imagine that you are an analyst and about to analyze the data for the first time.” After each animation, the participant had to answer *true*, *false,* or *can't tell* to a set of statements concerning spatial or temporal relations on the animated map (e.g., “Events took place in Africa.”). The research questions were answered by comparing the numbers of *correct*, *incorrect* and *can't tell* answers between the participants watching the animations with sound and those watching the animations without sound.

Each animation was followed by six statements presented in a random order; as the test included four different animations this made a total of 24 statements for each participant. The statements, along with their correct answers, can be found in *Table 2*. The goal was to make the statements as similar as possible between the animations with the same task variables. The statements refer to the detonations only by color, not by country.

	Statements	**Correct answer**
Animation 1 spatial-low	All detonations were north of the equator	FALSE
	Detonations took place in Africa	FALSE
	Blue detonations took place both in the sea and on land	TRUE
	The violet detonations all took place roughly in the same location	TRUE
	The detonations in North America were concentrated to the western parts	TRUE
	All detonations in Australia were violet	TRUE
Animation 2 temporal-low	The last violet detonation took place after the blue detonations	TRUE
	The beginning of January was detonation-free	TRUE
	Blue and violet detonations never took place at the same time	FALSE
	The violet detonations took place within an interval of two seconds	FALSE
	There were clearly more red detonations in the first half than the second half of the animation	FALSE
	The longest period with continuous detonations was longer than one second	TRUE
Animation 3 spatial-high	All detonations were north of the equator	TRUE
	Detonations took place in Africa	TRUE
	Green detonations took place both in Africa and in the Pacific Ocean	FALSE
	The violet detonations all took place roughly in the same location	TRUE
	The detonations in Africa were concentrated to the western parts	TRUE
	All detonations in North America were blue	FALSE
Animation 4 temporal-high	All yellow detonation took place before the green detonations	TRUE
	The beginning of January was detonation-free	TRUE
	Yellow and green detonations never took place at the same time	TRUE
	The yellow detonations took place within an interval of three seconds	FALSE
	There were clearly more blue detonations in the second half than in the first half of the animation	FALSE
	The longest period without any detonations was shorter than two seconds	TRUE

Table 2. The statements, along with their correct answers.

These types of tasks were chosen for two reasons. First, we wanted to simulate exploratory analysis, i.e., analysis where there is no clearly defined task but where the analyst tries to find patterns in the data. Therefore the statements were not given beforehand as this would have led the participants to focus only on what was mentioned in the statements. Second, we wanted a rigorous study design that would result in quantitatively analyzable data. A task where, e.g., the participants first watched the animated map and then listed all the insights they gained while watching it would have simulated exploratory analysis better, but it would have led to a less rigorous study design where the results would have been more difficult to analyze as each participant most probably would have come up with different numbers and types of insights. The drawback with this kind of design is that it adds difficulty to the task as the participants have to rely on their memory when answering the statements.

This type of task does not measure analytical reasoning directly but is a kind of simplification that can reflect how well animated maps facilitate exploratory analysis and pattern recognition. The statements used in the test were the kinds of insights that could be gained while watching the animations and asking the participants if they were true was one way of checking if they had gained the same kinds of insights.

There was also a fifth animation, of similar length and from the same artwork, that was followed by another type of task where the instructions were almost identical but where the participant had to state which kinds of events appeared in the animation and order them according to how many times they appeared. This task did not concern spatial or temporal aspects of the data but thematic ones and is therefore not analyzed in this paper. However, it might have had a modest influence on the participants' answers in the qualitative part of the test and it added about 3 minutes to the total duration of the test. Therefore it is briefly described below as a part of the test procedure.

2.3. Apparatus

The study was made using an internet-based survey system (Webropol[3]). However, because this survey system could not handle animations of the size needed in this study, the animations were uploaded to YouTube and then embedded in external web pages that included only one animation each. The YouTube player was configured in such a way that all the unnecessary and unwanted features were disabled, e.g., it was impossible to watch the animations in full screen mode. However, it was not possible to control, e.g., how many times the participants watched the animations or if they paused it or not.

---

[3] http://w3.webropol.com/int/

2.4. Test Procedure

After agreeing to take part in the study the participants got an individual link to the test in their e-mail. The test was named *Animations with spatial and temporal relations*. The test started with a general introduction where it was stated that the participants should be in a calm and quiet environment when performing the test. The number of animations included in the test and their length was stated, as was the estimated time for completing the whole test (a little less than 20 minutes). All the instructions were given in written form.

Then followed a more specific introduction to the animations, which included an introductory video. The video was a short annotated piece of the artwork explaining its basic features. This video, like the animations used later, opened in a new window or tab in the participants' browser. The participants were instructed to close it and return to the webropol page after watching it. Before the actual test started, the participants were also introduced to the map used in the animations and the most important geographical features on it (as some of them would be referred to in the statements), along with the legend that was shown next to the animations.

When the actual test was about to start the participants were urged to watch each animation thoroughly two times and not to go back to the animation after having read the statements. Moreover, the participants were told to answer "can't tell" if they did not know the answer instead of trying to guess.

In the actual test the participants were faced with the four animations described above one at a time and after each animation all six associated statements were shown together. The participants had to choose true, false or can't tell for each of the six statements before viewing the next animation. The fifth animation that was left out of this analysis was shown after this. The test was self-paced and the participants were not provided with feedback on any part of the test.

After this quantitative part the participants had to answer a couple of qualitative questions. First, the participants were asked for general comments ("Have you any comments about the animations? Was it easy to answer the questions or was it difficult?"), then there were four questions regarding the use of sound in map animations (both in general and in the animations used in the test). After this, there were some wrap-up questions about how the test had gone, including an explicit request to the participant to state if there had been any technical problems or if the instructions had not been followed correctly. The participants were also asked if the animations were familiar from before and if they had benefited from previous knowledge about the topic of the animations. Hence, it was possible to control for unsuccessful completions of the test and to check if anyone had benefited greatly by being too familiar with the animations or the theme of the animations. Finally some background information on age, sex, years of GIS experience and level of education was gathered and an Ishihara Color Vision Test was performed.

2.5. Participants

Participants were gathered through the e-mail lists of the Commission of Use and User Issues and the Commission on Cognitive Visualization of the International Cartographic Association and through the personal contacts of the first author. In total there were 24 participants in the study. These were randomly assigned to either the group watching the animations with sound or to the group watching the animations without sound. After both the quantitative and qualitative results had been gone through, four participants were left out of the analysis. Three of them stated they had not followed the instructions correctly and one performed remarkably worse than everyone else and was therefore considered as an outlier.

The remaining 20 participants happened to be equally divided between the two groups. Out of them, 10 were men and 9 women, and one did not state their sex. The youngest participant was 25 years old and the oldest was 77 years old; the median age was 32.5. There were no decisive differences between the groups when it came to sex and age distribution. All the participants except one had a Master's or Doctoral degree and all except one had three or more years of GIS experience. The participants were tested for colorblindness with the Ishihara Color Vision Test and everyone passed it. The participants did not get any compensation for taking part in the test.

3. Results

The results of the quantitative and the qualitative parts of the test are analyzed separately. *Section 3.1* accounts for the effect of the audio variable and *Section 3.2* reviews the qualitative part of the test.

3.1. Answers to the statements

In total there were 480 answers (2 groups, 10 people/group, 4 animations/person, 6 statements/animation). The answers can be found in *Table 3*, along with the differences induced by the audio variable in percentage points. The differences are relatively small, with 13 percentage points being the largest (temporal-low). Otherwise, the most general trend is that the spatial statements yielded a higher portion correct answers than the temporal statements (83 % vs. 21 % without sound and 83 % vs. 26 % with sound). The temporal statements, on the other hand, yielded a higher portion "can't tell" answers (68 % vs. 6 % without sound and 62 % vs. 8 % with sound). The event frequency variable did not affect the answer rates as much as the type of statement did (the task variable).

Statements	No sound				Sound				Difference (sound – no sound)		
	n	Answers [%]			n	Answers [%]			Percentage points		
		Correct	Wrong	Can´t tell		Correct	Wrong	Can´t tell	Correct	Wrong	Can´t tell
All	240	52	11	37	240	54	11	35	2	0	-2
All spatial	120	83	11	6	120	83	9	8	-1	-2	3
All temporal	120	21	11	68	120	26	13	62	5	2	-7
Spatial-low	60	90	3	7	60	97	2	2	7	-2	-5
Spatial-high	60	77	18	5	60	68	17	15	-8	-2	10
Temporal-low	60	20	7	73	60	20	20	60	0	13	-13
Temporal-high	60	22	15	63	60	32	5	63	10	-10	0

Table 3. The results of the user test. The difference induced by the audio variable is shown as percentage points: e.g., a difference in correct answers of 2 means that there were 2 percent more correct answers in the group that saw the animations with sound than in the group that saw it without sound. In the Correct column, a *positive* value indicates a positive effect of adding sound (more correct answers), while in the Wrong and Can't tell columns, a *negative* value indicates a positive effect of adding sound (fewer wrong and can't tell answers).

The audio-variable results were tested for statistical significance in all possible sets. A χ^2 test of independence was used with a p-value of .05. In the cases where the minimum expected value was less than 5 a Fisher exact test with a p-value of .05 was used instead. The results can be found in *Table 4*. For information about the tests, see, e.g., McDonald (2014).

Statements	χ^2 test		Fisher exact test
	χ^2	p-value	p-value
All	0.24	.89	-
Spatial	0.70	.70	-
Temporal	1.20	.55	-
Spatial-low	-	-	.40
Spatial-high	3.34	.19	-
Temporal-low	4.80	.09	-
Temporal-high	4.12	.13	-

Table 4. The results of the test of statistical significance.

No statistically proven effect of the audio variable was found. The subset that came closest to the p-value of .05 was that of temporal-low statements, with a χ^2 of 4.80 and a p-value of .09.

3.2. Opinions about the Animations

When asked for general comments about the test, 17 out of 17 answers mentioned that the test was difficult in some way and 10 mentioned that especially the temporal statements were difficult. The questions about the use of sound differed somewhat between the two groups as one had seen the animations with sound and knew what it meant, while the other group, which had not viewed the animations with sound, had to imagine what it would be like. (The first group were only offered *yes* or *no* as alternatives, the second group could also answer *can't tell*.)

When asked about whether sound made it easier to take in information about the time, the place and the origin of the detonations, and whether it was especially beneficial for any of these in particular, the answers were similar between the two groups. A total of 13 out of 20 answered that sound made it easier, or would have made it easier, and 14 thought it would be helpful for temporal relations, 2 for spatial relations and 4 for thematic information, that is, the country performing the tests. When asked if sound made it easier, or would have made it easier, to answer the statements, the answers differed between the two groups. Of those who viewed the animations without sound 8 out of 10 answered yes and out of those who watched it with sound only 4 out of 10 thought it was helpful.

Finally, the participants were given a chance to say in their own words what they thought about the possibility of using sound in map animations of this kind. Out of 16 answers, 11 thought it would have a positive effect and 2 thought it would have a negative effect. Those that thought it would be negative mentioned an increased cognitive load and even that "it would probably drive people mad..."[4] Four participants mentioned that sound should be used sparsely or carefully, e.g., "If sounds were used, only one information parameter should represented using audio aid [sic]."

[4] This person saw the animation without sound.

4. Discussion

In this section, the results presented in the previous section are discussed. Some interesting trends in the results that have not previously been mentioned but need to be taken into account in future research are also introduced, along with some general recommendations for future research.

4.1. Sound versus No Sound

No statistically proven effect of the audio enhancement of map animations on communicating spatial and temporal relations in an exploratory analysis context could be seen in the results. However, it seems that added sound does not impair the perception of these relations either, even when they encode redundant information, as in this case (i.e., the sound did not add any new information to the animations). Probably, Tufte's principles (2001) about the data-ink ratio do not apply in multimodal settings. Instead, users are most likely to benefit when the same information is being presented in multiple modalities, as we humans benefit from the redundancies provided to us in real life (Alais et al. 2010). Auditory perception does not affect visual perception negatively, and vice versa.

Despite the results the authors of this paper still hold the belief that the hypothesis might be true. The opinions of the participants also support this belief: 13 out of 20 thought sound made it easier to take in the information in the animations, and 14 out of 20 thought it was helpful for perceiving temporal relations. When it came to the statements, the participants made the same error as the authors: those that saw the animations without sounds thought sound would be helpful (8 out of 10), while those that saw it with sound knew better (only 4 out of 10 thought it was helpful). However, these numbers should be considered with some caution as user preferences not always correlate with efficient performance (Hegarty et al. 2009).

That the hypothesis could not be confirmed in this study might depend on an inappropriate study design, and not on the hypothesis being wrong. Maybe the tasks used were too great a simplification of what analytical reasoning really is or simply too difficult to reveal any differences between the two settings. Additionally, the task taxonomy, dividing the statements into spatial and temporal, might have been too general. In a way, this study confirms what Fuhrmann et al. (2005) said about measuring usefulness in an exploratory analysis context: it is difficult.

It might also be that the animations used in this study were not that well suited to being used as stimuli. Here, it is important to bear in mind the original purpose of the artwork: Hashimoto wanted to attract attention to what he calls an “extremely grave, but present problem of the world” (CTBTO Preparatory Commission 2012). He wanted to provoke an emotional reaction in the viewer. The extent to which the artwork has been shared on the internet can be seen as proof of him succeeding in that respect. The sounds are an important part of this as they carry emotional information (Weninger et al. 2013), or, as one participant phrased it, they make it more “dramatic”. The artwork was not designed to facilitate analytical reasoning. It is possible that we, the authors, were also emotionally affected by the sounds in the artwork, and deceived ourselves into thinking that they facilitated our analytical reasoning.

Probably, sounds should be used with reflection and implemented on a case-by-case basis. The fact that it is not possible to ignore audio stimuli in the same way as visual stimuli can be seen as both an advantage and disadvantage. Visual perception requires that you fix your eyes upon the target while sounds are heard regardless of in what direction you are facing. Therefore, it is not necessary to consciously attend to sonifications in the same way as to visualizations, but then again, irritating sounds can be very frustrating. Several participants also made their concern about this known in the test, mentioning, e.g., that bad sound implementation would have been irritating and confusing.

4.2. Observations for Future Research

The results include a lot of interesting trends that, although not statistically significant, deserve to be mentioned, and that need to be taken into account in future research. Furthermore, the statistical significance of the *event frequency* variable and the *task* variable was tested using the same procedure as above. As the test was not originally designed for this kind of analysis it might not produce as rigid results for these variables as for *audio*; however, the individual observations should be independent enough for this type of analysis as well, as the animations were different for each set of statements. To keep the length of the article down, this part does not present all the exact numerical values.

One interesting trend can be seen when comparing the effect of the sound variable statement by statement between the high and low event frequencies (i.e. comparing statements 1-6 in spatial-low pairwise with statements 1-6 in spatial-high and likewise for temporal-low and temporal-high). It seems that for the spatial task, the added sound had a more positive influence on performance when there were few events compared with when there were many events. For the temporal task, the trend was the opposite; here the added sound had a more positive influence when there were many events.

This raises several questions. Did the sounds function as attention grabbers (minimizing change blindness) in the spatial-low animation and did this positive effect disappear in the spatial-high animation? Did the sound function as a temporal reference frame, making it easier to relate the different events temporally to each other in the temporal-high animation? Did this effect not take place in the temporal-low animation? These are questions that have to be addressed by future research.

The event frequency variable also had a different effect on the participants' performance on the spatial and the temporal task. The participants performed better on the spatial-low animation than the spatial-high animation, regardless of whether the animation was accompanied by sound or not, which was quite expected. If there are few things happening, they are easier to perceive and remember compared with if there are many things happening. Statistical testing confirms these results. The difference induced by the event frequency is statistically relevant for both the set of all spatial statements (χ^2 test: $\chi^2 = 19.52$, $p < .01$) and for spatial-sound (χ^2 test: $\chi^2 = 16.68$, $p < .01$) and spatial-no sound (Fisher exact test: $p = .04$).

However, there is no such clear trend for the temporal task. The only statistically provable difference in performance between high and low event frequencies is found for temporal-sound statements (χ^2 test: $\chi^2 = 7.04$, $p = .03$) and there the high event frequency yielded better performance, i.e., the opposite to the case of the spatial task. This seems to strengthen the previously mentioned trend: for the temporal task a positive effect of sound can mainly be found for animations with high event frequency. Still, it seems counterintuitive that the statements were easier to answer when there were many events compared with when there were few events. Future research needs to address these matters and to try to figure out the exact cause and effect relationships so that the possibility that it was not simply due to chance differences in the stimuli and statements can be excluded.

Overall, the spatial statements were easier to answer than the temporal statements (χ^2 test: $\chi^2 = 192.18$, $p < .01$). Definitely, one reason for this was that it was easier to locate, and hence remember, the events in relation to the spatial frame of reference, i.e., the world map, than it was to locate them in relation to the temporal frame of reference, i.e., the calendric system. The world map was always visible in the background, while the temporal reference frame was communicated through numbers on the periphery and through only two different kinds of sounds. Previous research has already pointed to the important role of the reference system in analytical reasoning of this kind (Hall & Ahonen-Rainio 2014). Other reasons are probably that we have no specific biological system that senses time (Merchant et al. 2013), while vision works excellently for sensing place, and that reasoning about space is one of the most common and basic forms of human intelligence (Egenhofer & Mark 1995, Levinson 2003).

Future research should use different tasks and be done with animations especially designed to test if sounds facilitate exploratory analysis and analytical reasoning. To answer the questions in this section future research needs to apply a design that allows the identification and isolation of the exact cause and effect relationships. This requires animations that apply sound for one thing at a time, minutely thought-through tasks and more versatile measurement variables. One possibility is to measure the time taken to complete a task in addition to the accuracy of the answer. The use of spatialized sound is also an interesting possibility for future research.

5. Conclusions

Many authors have pointed to the need for multimodal approaches to interface design in geovisualization, and particularly the use of sound to enhance map animations has been brought up several times. However, practical experiences of using sound in map animation are hard to find in the literature. This study, using animations that applied sound as a temporal legend and to accompany events, was a reaction to this deficit but failed to deliver an unambiguous answer.

What can be said on the basis of this study is that adding sound does not impair the usability of an animated map in a visual analytics context, but it is not possible to say that it improves it either. It might be that the experts who have commented on this subject in the literature, as well as the authors of this paper and the participants in the study, are all wrong in thinking that sound can be used to enhance map animations in an exploratory analysis context, but a more reliable conclusion is that this study was not appropriately designed to measure what it set out to measure.

The test produced some interesting results indicating that the benefit of sound in animated maps might be related to whether there are many or few events in the animation. There also seems to be a difference in the benefit of adding sound when it comes to reasoning about temporal and spatial relations. The results also indicated that reasoning about temporal relations might be easier when there are many events, a result opposite to that of reasoning about spatial relations. No definite conclusions can be drawn on the basis of this study but it works as a good starting point for future research.

References

Aigner W, Miksch S, Schumann H et al (2011) Visualization of time-oriented data. Springer

Alais D, Newell FN, Mamassian P (2010) Multisensory Processing in Review: from Physiology to Behaviour. Seeing and Perceiving 23(1):3-38

Andrienko G, Andrienko N, Demsar U et al (2010) Space, time and visual analytics. International Journal of Geographical Information Science 24(10):1577-1600

Bearman N, Lovett A (2010) Using sound to represent positional accuracy of address locations. The Cartographic Journal 47(4):308-314

Bertin J (1983) Semiology of graphics. Diagrams, Networks, Maps. University of Wisconsin press, 1983 (first published in French in 1967, translated to English by Berg WJ in 1983)

Bolognini N, Cecchetto C, Geraci C et al (2012) Hearing shapes our perception of time: temporal discrimination of tactile stimuli in deaf people. Journal of Cognitive Neuroscience 24(2):276-286.

CTBTO Preparatory Commission (2012) "1945-1998" by Isao Hashimoto. http://www.ctbto.org/specials/1945-1998-by-isao-hashimoto/. Accessed 2 October 2014

Dubus G, Bresin R (2013) A systematic review of mapping strategies for the sonification of physical quantities. PLOS One 8(12):e82491. doi: 10.1371/journal.pone.0096018

Egenhofer M, Mark D (1995) Naive geography. Spatial Information Theory: A Theoretical Basis for Gis 988:1-15

Fairbairn D, Andrienko G, Andrienko N et al (2001) Representation and its relationship with cartographic visualization. Cartography and Geographic Information Science 28(1):13-28

Fisher P (1994) Animation and sound for the visualization of uncertain spatial information. In: Hearnshaw HM, Unwin DJ (eds) Visualization in Geographical Information Systems, p 181-185. Wiley & Sons

Fröhlich B, Barrass S, Zehner B et al (1999) Exploring geo-scientific data in virtual environments. In: Proceedings of Visualization'99, p 169-173. IEEE

Fuhrmann S, Ahonen-Rainio P, Edsall RM et al (2005) Making useful and useable geovisualization: design and evaluation issues. In: Dykes J, MacEachren AM, Kraak M-J (eds) Exploring Geovisualization, p 553-566. Pergamon

Griffin AL, MacEachren AM, Hardisty F et al (2006) 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):740-753

Hall A, Ahonen-Rainio P (2014) Analysis of Basic Relations within Insights of Spatio-temporal Analysis. In: Vondrakova A, Brus J, Vozenilek V (eds) Modern Trends in Cartography, Selected Papers of CARTOCON 2014, p 409-423. Springer

Harding C, Kakadiaris IA, Casey JF et al (2002) A multi-sensory system for the investigation of geoscientific data. Computers & Graphics 26(2):259-269

Harrower M (2007) The cognitive limits of animated maps. Cartographica: The International Journal for Geographic Information and Geovisualization 42(4):349-357

Harrower M, Fabrikant S (2008) The role of map animation for geographic visualization. In: Dodge M, McDerby M, Turner M (eds) Geographic Visualization: Concepts, Tools and Applications, p 49-65. Wiley

Hegarty M, Smallman HS, Stull AT et al (2009) Naïve cartography: How intuitions about display configuration can hurt performance. Cartographica: The International Journal for Geographic Information and Geovisualization 44(3):171-186

Heming JE, Brown LN (2005) Sensory temporal processing in adults with early hearing loss. Brain and Cognition 59(2):173-182

Kowalska J, Szelag E (2006) The effect of congenital deafness on duration judgment. Journal of Child Psychology and Psychiatry 47(9):946-953

Kraak M, Edsall R, MacEachren AM (1997) Cartographic animation and legends for temporal maps: Exploration and or interaction. In: Proceedings of the 18th International Cartographic Conference

Kramer G, Walker B, Bargar R (1999) Sonification report: Status of the field and research agenda. International Community for Auditory Display

Kriglstein S, Pohl M, Smuc M (2014) Pep Up Your Time Machine: Recommendations for the Design of Information Visualizations of Time-Dependent Data. In: Huang W (ed.) Handbook of Human Centric Visualization, p 203-225. Springer

Krygier JB (1994) Sound and Geographic Visualization. In: MacEachren AM, Taylor DRF (eds) Visualization in Modern Cartography, p 149-166. Pergamon

Laakso M, Sarjakoski TL (2010) Sonic Maps for Hiking—Use of Sound in Enhancing the Map Use Experience. The Cartographic Journal 47(4):300-307

Lam H, Bertini E, Isenberg P et al (2012) Empirical studies in information visualization: Seven scenarios. IEEE Transactions on Visualization and Computer Graphics 18(9):1520-1536

Levinson SC (2003) Space in language and cognition: Explorations in cognitive diversity. Cambridge University Press

MacEachren AM, Kraak M (2001) Research challenges in geovisualization. Cartography and Geographic Information Science 28(1):3-12

Mayer RE (2005) Cognitive Theory of Multimedia Learning. In: Mayer RE (ed) The Cambridge Handbook of Multimedia Learning, p 31-48. Cambridge University Press

McDonald JH (2014) Handbook of biological statistics. Sparky House Publishing

Merchant H, Harrington DL, Meck WH (2013) Neural basis of the perception and estimation of time. Annual Review of Neuroscience 36:313-336

Midtbø T (2001) Visualization of the temporal dimension in multimedia presentations of spatial phenomena. In: Bjørke JT, Tveite H (eds).

ScanGIS'2001, Proceedings of the 8th Scandinavian Research Conference on Geographical Information Science, Ås, Norway, p 213-224

Reuter LH, Tukey P, Maloney LT et al (1990) Human perception and visualization. In: Proceedings of the 1st conference on Visualization'90, p 401-406. IEEE Computer Society Press

Simons DJ, Rensink RA (2005) Change blindness: Past, present, and future. Trends in Cognitive Sciences 9(1):16-20

Slocum TA, Blok C, Jiang B et al (2001) Cognitive and usability issues in geovisualization. Cartography and Geographic Information Science 28(1):61-75

Tufte ER (2001) The visual display of quantitative information, 2nd edition. Graphics Press

Tversky B, Morrison JB, Betrancourt M (2002) Animation: can it facilitate? International Journal of Human-Computer Studies 57(4):247-262

Weninger F, Eyben F, Schuller BW et al (2013) On the acoustics of emotion in audio: what speech, music, and sound have in common. Frontiers in Psychology 4

Zhao H, Plaisant C, Shneiderman B et al (2008) Data sonification for users with visual impairment: a case study with georeferenced data. ACM Transactions on Computer-Human Interaction (TOCHI) 15(1):4