

# Both Rudimentary Visualization and Prototypical Sonification can Serve as a Benchmark to Evaluate New Sonification Designs

TIM ZIEMER and HOLGER SCHULTHEIS\*, Bremen Spatial Cognition Center, University of Bremen, Germany

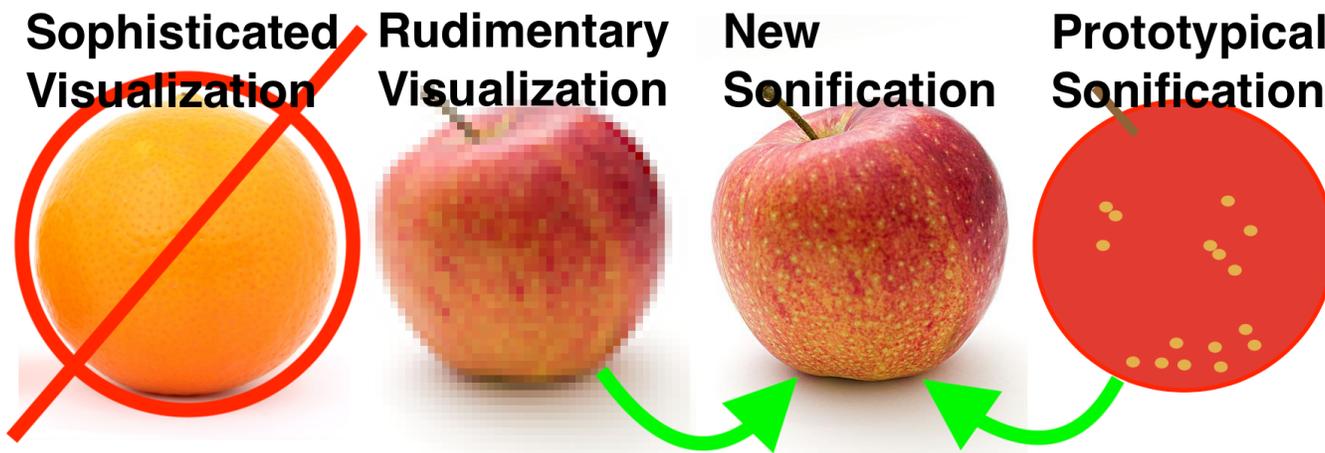


Fig. 1. Instead of comparing apples and oranges, rudimentary visualization and prototypical sonification can serve as a benchmark for the evaluation of new sonification designs.

Comparing sonification with visualization is like comparing apples and oranges. While visualizations are ubiquitous to the public and have established names, principles, application areas, and sophisticated designs, sonifications tend to be unique, self-made and completely new to users. In this study we developed a rudimentary visualization that is related closely to the principle of the sonification designs that we want to evaluate. In addition, we implemented a prototypical sonification that uses the most common mapping principles. Experiment results show that participants perform similarly well using the rudimentary visualization and the prototypical sonification, which is much better than chance but significantly worse than using our new sonification results. We therefore argue that both rudimentary visualization and prototypical sonifications can serve as a suitable benchmark to evaluate new sonifications designs against.

CCS Concepts: • **Human-centered computing** → *Visualization techniques*;  
• **Computer systems organization** → **Embedded systems**;

Additional Key Words and Phrases: sonification evaluation, auditory display evaluation, visualization

## 1 INTRODUCTION

Sonification is a comparably young discipline of design practice and research [13, Chap. 1]. A review study by [3] revealed that only a marginal proportion of studies report a proper evaluation of their sonification.

Studies like [1, 2, 8, 15] evaluate their sonification design against sophisticated visualizations in (data) navigation guidance tasks. The typical result is that users perform better using visualization rather than sonification, especially in terms of the time needed to reach a target. Naturally, this comparison is unfair because participants tend to have life long experience and everyday practice using the respective visualization for navigation, from traffic signs over automotive navigation systems, maps and compasses, to graphical user interfaces in video games and routing apps. Compared to that many people have never even heard of sonification, let alone consciously used it. The comparison of sophisticated and well-established visualization with new sonification designs is like a comparison of apples with oranges indicated in Figure 1. In this example the visualization is an orange and the sonification to be evaluated is the photo of the apple. Results from such a comparison should be analyzed with care.

In the study at hand we aimed at designing a rudimentary visualization that can serve as a benchmark for a fair comparison between sonification and visualization. This rudimentary visualization is not a sophisticated and well-established visualization, but a visualization that is closely related to the principles that the evaluated sonifications use. The rudimentary sonification is indicated as a pixelated apple in Figure 1. It can serve as a more suitable benchmark.

[10] compare their four novel sonification designs with a number of sonifications that are based on pitch, loudness, duration and brightness. As these sonification designs are among the most frequently used parameter mappings [3], these can be considered as *prototypical sonifications* that are well-suited benchmarks. The sonifications informed about the absolute distance between the user and an invisible target located somewhere on the right hand side. When the task was to find the target with the highest possible precision, users were significantly more precise using two of their sonification designs, compared to all other sonification designs. When the task was to find the target as quickly as possible, one of their sonification designs was significantly more precise, but none was significantly faster (or slower). Other studies, like [4, 7, 9], use a similar approach, comparing novel sonification designs with prototypical sonifications. In Figure 1 the computer-drawn apple represents a prototypical sonification. It is similar to the real apple, yet not as advanced.

In the study at hand we use a prototypical sonification not only as a benchmark to evaluate our own sonification designs, but also as a means to validate how well our rudimentary visualization is suited as a benchmark to evaluate our sonification designs.

The remainder of the paper is structured as follows: Section 2 explains the method of our experiment in which participants had to assign a presented data point to the respective field on a map. It starts with the participants' task, followed by the explanation that we gave them, and the data presentations methods that we used. These methods are our rudimentary visualization design that is supposed to present these data points in a way that is comparable to sonification, our own sonification designs that we want to evaluate, and a prototypical sonification that is a suitable benchmark. The section closes with an introduction of the measures that we use to analyze our experiment results.

Section 3 presents the performance of the participants using the visualization or either of the sonifications. In Section 4 we discuss to what extent the results indicate the appropriateness of the rudimentary visualization and the prototypical sonification as a benchmark to evaluate new sonification designs against.

## 2 METHOD

We carried out an experiment that took about 10 minutes. We randomly assigned 35 participants (mostly computer science students and researchers, aged 20 to 53, mean 27, 4 females and 31 males) to one out of 5 groups. Each group had a different condition, i.e., a different data presentation method.

### 2.1 Task

First, we explained to the participants that we were dealing with a two-dimensional Cartesian coordinate system in which each dimension ranged from  $-1$  to  $1$  as illustrated by the arrows in Fig. 2. We further explained that we subdivided this coordinate system into 16 fields as seen in Fig. 2. We told the participant that a sequence of coordinate pairs would be presented to them, each representing a different positions within this coordinate system. Their task was to assign each of the 20 presented positions to one of the fields and mark it on a map that was printed on a paper sheet. We told them

that they should mark one field for each presented position, that they would not receive any feedback whether their assignment was correct or not, and we told them not to amend their assignment. We informed the participants that the data points would be played in random order, that the sequence would not necessarily cover all 16 fields, and that fields could occur multiple times. This way we assured that participants would not deduce the respective fields for the last few data points as those fields that have not been occupied before. Indeed, for each participant at least one data point for each field was presented.

### 2.2 Preparation

We explained the data presentation method to the participants, supported by the graphics as shown in Figures 2 to 5. We explained how the  $x$ - and  $y$ -coordinates would be represented through their respective visualization or sonification method. All presentation methods followed the same principle: each direction (polarity) of both dimensions is represented by its individual characteristic (either of a graphic or a sound). The intensity of the characteristic indicates the distance along that direction. In addition to the explanation we presented the participants a series of coordinates along the axes, namely  $(-0.8, 0)$ ,  $(-0.75, 0)$ ,  $(-0.3, 0)$ ,  $(0.3, 0)$ ,  $(0.75, 0)$ ,  $(0.8, 0)$ ,  $(0, -0.8)$ ,  $(0, -0.75)$ ,  $(0, -0.3)$ ,  $(0, 0.3)$ ,  $(0, 0.75)$ ,  $(0, 0.8)$ . We told the participants that all data points would be combinations of the respective  $x$ - and  $y$ -component throughout the experiment. However, we did not present combinations to them. They would have to understand the mapping principle, infer how combinations should look or sound like, and remember the principle throughout the experiment. The explanation took about 5 minutes.

To group 1 we presented the data points by means of visualization, to groups 2–4 we presented the data points by means of our own sonification designs, and to group 5 we presented the data points by means of a prototypical sonification. The data presentation methods are described below.

### 2.3 The Visualization

The visualization is based on a simple color space. It is illustrated in Fig. 2. The coordinate pair is represented by a colored dot as demonstrated on <https://youtu.be/QZ3PEyTi6sI>.

The  $x$ -dimension is the red/blue dimension. From left to right the intensity of the red color component reduces linearly. Likewise, the percentage of the blue color component increases linearly from left to right. Consequently, an  $x$ -value of 0 looks violet.

The  $y$ -dimension is divided in two. The positive  $y$ -polarity is the yellow polarity. The larger the  $y$ -value, the higher the percentage of the yellow component. Note that according to the RGB color scheme this produces orange in the upper left corner, turquoise in the upper right corner, etc. The negative  $y$ -polarity is the brightness polarity. The larger the negative value, the darker the color. This produces dark red in the lower left corner, dark blue in the lower right corner, etc.

This rudimentary visualization is supposed to serve as a benchmark. It presents the data in a way that is related to the sonifications:

- (1) The visualization uses a mapping scheme
- (2) Each polarity is represented by a certain characteristic.

- (3) The intensity of this characteristic represents the magnitude of the respective polarity, i.e., the distance along that direction.
- (4) No references in terms of the map, coordinate axes or boundaries is provided during the data presentation.
- (5) Instead, the data presentation method has to be understood and interpreted during the trials.
- (6) The sequence of data points gives additional clues. For example, if the intensity of a certain characteristic seems larger than that of the previous data point, this indicates that the associated field lies further along the respective direction that is represented by this characteristic.
- (7) Moreover, participants are certainly not familiar with the concept of representing positions by means of color and brightness (i.e., the visualization method is atypical compared to well-established and ubiquitous visualizations)

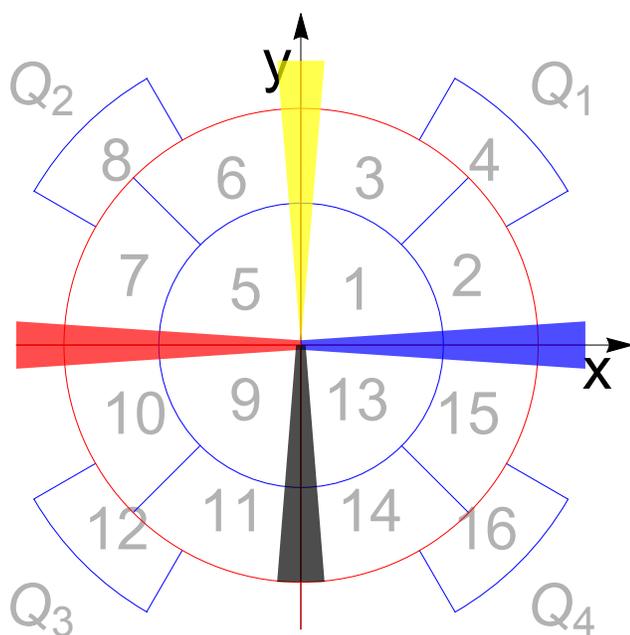


Fig. 2. Symbolic indication of the visualization (foreground) which indicates data points. The data points need to be assigned to the respective field on the map shown in the background. The map has 16 fields in its 4 quadrants (Q).

#### 2.4 Chroma vs. Beats/Roughness Sonification

This sonification is based on a Shepard tone [12]. 12 partials in octave intervals have a frequency-dependent amplitude envelope that is tapered off towards the ends of the audible frequency range.

The  $x$ -dimension is the *chroma*-dimension  $x$ . The  $x$ -coordinate is mapped to the frequency of a phasor that sweeps all 12 frequencies exponentially through the range from  $2^2$  to  $2^{14}$  Hz. Positive  $x$ -coordinates sweep them upwards. This sounds like a rising pitch, while in fact the cyclic chroma moves clockwise. As soon as a frequency has reached  $2^{14}$  Hz, it is reset to  $2^2$  Hz. This creates the

illusion of an ever-rising pitch, known as *Shepard illusion*. Negative  $x$ -coordinates simply sweep the other way. The larger the magnitude of  $x$ , the faster the sweep. Consequently, a small positive  $x$ -value sounds like a slowly rising pitch (like increasing the string-tension while tuning a guitar), a large  $x$ -value sounds like a quickly rising pitch (here, you start hearing the cyclic frequency motion, so the result sounds a bit siren-like). Likewise, negative  $x$ -values sound like a falling pitch. Only at  $x = 0$  the pitch is steady. The principle is indicated symbolically in Fig. 3, together with the  $y$ -dimension and the map. A demonstration of the series of coordinates can be found on <https://youtu.be/4hlfPoYfVrg>.

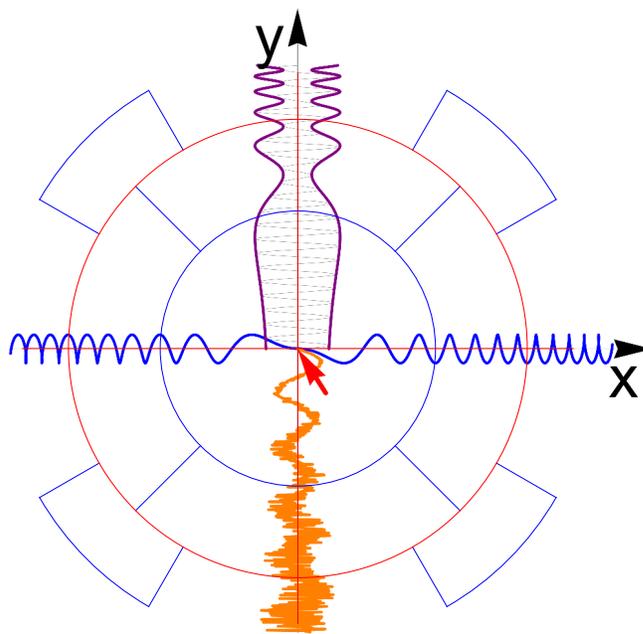


Fig. 3. Symbolic visualization of the chroma vs. beats/roughness sonification over the map.

The  $y$ -dimension is divided in two. The positive  $y$ -polarity is the *beats* polarity. Here, the gain of the Shepard tone is modulated using a low frequency oscillator whose frequency is proportional to  $y$ . This sounds like loudness fluctuation or beats. The larger the  $y$ -value the faster the beats. The negative  $y$ -polarity is the *roughness* polarity. All 12 frequencies of the Shepard tone served as carrier frequencies in a frequency modulation synthesis. The modulation frequency is 50 Hz, creating sidebands near the carrier frequency, which produce a rough sound. The magnitude of the  $y$ -value is mapped to the modulation index. The larger the magnitude of  $y$ , the higher the number and amplitude of sidebands, i.e., the rougher the resulting sound.

In the upper left corner, the sound exhibits both a quickly falling pitch and fast beats. In the upper right corner the sound exhibits a quickly rising pitch and fast beats. In the lower left corner, a very rough sound with quickly falling pitch is audible, etc.

## 2.5 Chroma vs. Fullness/Brightness Sonification

Again, the core of this sonification is the Shepard tone, and the  $x$ -dimension equals the previous  $x$ -dimension.

Again, the  $y$ -dimension is divided in two. The positive  $y$ -polarity is the *fullness* polarity. Here, the  $y$ -value controls the height and width of the frequency-dependent envelope, affecting the physical bandwidth and the perceived fullness. At low  $y$ -values the spectrum covers the complete audible frequency range and is therefore perceived as full. At the maximum  $y$ -value the envelope is narrow and the sound is dominated by a single frequency, sounding not full but empty or thin. Simply reducing the bandwidth reduces not only the perceived fullness, but also the perceived loudness. To counter balance this effect, the height of the envelope increased with decreasing bandwidth, yielding a constant loudness. In the upper left corner, a cyclically, quickly falling pitch with a pure-tone-like timbre is audible. The negative  $y$ -polarity is the *brightness* polarity. Here, the envelope is shifted towards higher frequencies for larger absolute values of  $y$ . This increases the spectral centroid and thus the perceived brightness. At  $y = 0$  the sound is a bit dull, whereas at  $y = -1$  the sound is shrill. The principle is illustrated in Figure 4 and demonstrated on <https://youtu.be/cW6EeBfEZp8>.

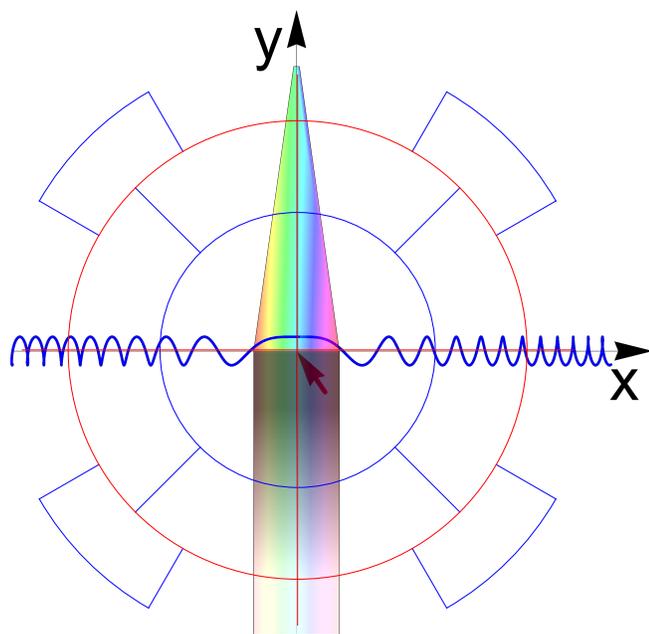


Fig. 4. Symbolic visualization of the chroma vs. fullness/brightness sonification.

## 2.6 Fullness/Brightness vs. Beats/Roughness Sonification

Again, the core of this sonification is the Shepard tone with its 12 carrier frequencies.

The  $y$ -dimensions of the two previous sonification designs are combined. This time the  $x$ -dimension is made of the fullness and brightness polarity and the  $y$ -dimension of the beats and roughness

polarity as illustrated in Fig. 5 and demonstrated on <https://youtu.be/ksr9eNTv7Tc>.

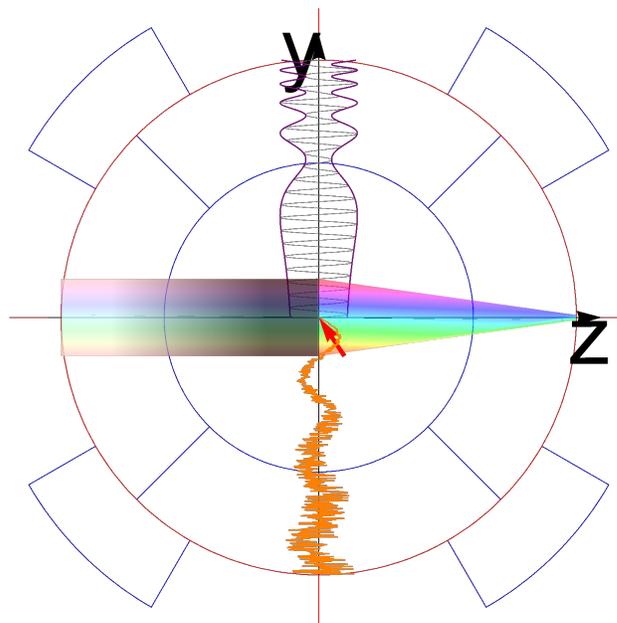


Fig. 5. Symbolic visualization of the fullness/brightness vs. beats/roughness sonification.

## 2.7 Panning vs. Frequency/Amplitude Sonification

This sonification is a prototypical sonification that is supposed to serve as a benchmark. It is based on the most frequent mappings according to the review study by [3]. The core of the sonification is a pure tone. The sonification principle is illustrated in Fig. 6 and demonstrated on <https://youtu.be/QT2nE5COFj8>.

The  $x$ -dimension is represented by amplitude based panning. Positive  $x$ -values produce a higher amplitude on the right loudspeaker, negative values produce a higher amplitude on the left loudspeaker, according to the tangent panning law.

The  $y$ -dimension is divided in two. The positive polarity is the *frequency* polarity. The higher the  $y$ -value the higher the frequency of the pure tone. The negative polarity is the *amplitude* polarity. The higher the  $y$ -magnitude, i.e., the lower the field on the map, the lower the amplitude of the pure tone. A logarithmic mapping for frequency and amplitude is chosen to produce a fair approximation between physical frequency and perceptual pitch, and between physical amplitude and perceptual loudness [5, pp. 134ff][11].

This prototypical sonification uses the sample principles as our own sonification designs: Each polarity of both dimensions is represented by its individual sound characteristic. The distance along that direction is represented by the intensity of the respective sound characteristic.

Such prototypical mappings have proven their value. The respective sound characteristics are quite fundamental and easy to recognize. But their limitations are known (cf. [6, pp. 34 & 40] [14]): they

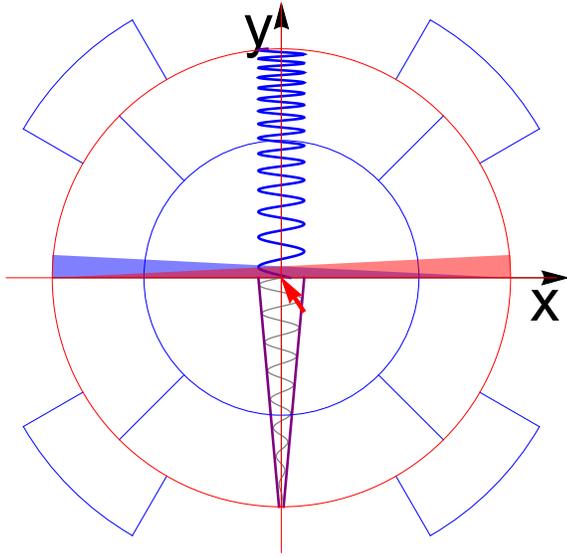


Fig. 6. Symbolic visualization of the panning vs. frequency/amplitude sonification.

tend to have a low perceptual resolution, lack absolute values, and potentially interfere perceptually with each other. Overcoming these limitations is one of the main reasons why new sonification designs are being developed. To evaluate the success of a new sonification design it is therefore meaningful to let a prototypical sonification act as a benchmark.

### 2.8 Measures

In all presentation methods each quadrant has its individual polarity pair, i.e., combination of visual or sound characteristics. To correctly recognize the presented field, it is important to identify these two respective characteristics. If successful, the participant will assign the presented data point at least to the correct quadrant. This is the minimum requirement of the data presentation method. To assign the presented data point even to the correct field, the intensity of each characteristic, i.e., the magnitude of each polarity, has to be identified successfully. This is the ultimate objective of the data presentation method.

We therefore evaluate two different measures: The frequency of correctly assigned quadrants ( $Q$ ), i.e., the number of data points that have either been assigned to the correct field or to one of the other 3 fields within the respective quadrant. And the hit rate ( $H$ ), i.e., the frequency of correctly assigned fields.

## 3 RESULTS

The experiment results are summarized in Table 1. It can be seen that the participants in our experiment performed similarly well using our three sonification designs, and better than when using the rudimentary visualization or the prototypical sonification. In all cases  $Q$  and  $H$  are much higher than expected by a random

method	$Q$			$H$		
	mean	$\sigma$	95% CI	mean	$\sigma$	95% CI
$v_{\text{ref}}$	64%	22%	43.3–83.7%	29%	15%	16.4–43.6%
S1	89%	12%	78.1–100%	60%	26%	36.1–83.9%
S2	91%	9%	83.8–99.3%	66%	22%	45.3–86.1%
S3	90%	10%	81.15–98.85%	51%	7%	22.4–58.35%
$S_{\text{ref}}$	68%	12%	56.9–78.8%	36%	15%	19.9–50.1%

Table 1. Percentage of correctly identified quadrants ( $Q$ ) and correctly assigned fields (“hits”  $H$ ), standard deviation ( $\sigma$ ) and 95% confidence interval (CI) for the rudimentary visualization  $v_{\text{ref}}$ , our three sonification designs (S1 to S3) and the prototypical sonification  $S_{\text{ref}}$ .

guess, which would lie at  $Q_{\text{guess}} = 1/4 = 25\%$  for  $Q$  and  $H_{\text{guess}} = 1/16 \approx 6\%$  for  $H$ . A benchmark should represent a baseline that is comparable to the state-of-the-art. As such, a minimum requirement for a benchmark is that it lies above chance level.

A one-way MANOVA reveals a significant dependence ( $F(8, 58) = 3.42, p < 0.05$ , Wilk’s  $\Lambda = 0.461$ , partial  $\eta^2 = 0.321$ ) of  $Q$  and/or  $H$  on the data presentation method. The tests of between-subjects effects reveals that (after Bonferroni correction  $\alpha_{\text{bon}} = \frac{\alpha}{2} = 0.025$ ) the presentation method had a significant effect on  $Q$  ( $F(4, 30) = 6.954, p < 0.0005$ , partial  $\eta^2 = 0.481$ ) and on  $H$  ( $F(4, 30) = 4.970, p < 0.005$ , partial  $\eta^2 = 0.399$ ).

Tukey post-hoc test reveals that our three sonification designs do not produce significant performance differences ( $p > 0.998$  for  $Q$  and  $p > 0.905$  for  $H$ ). In addition to the statistical significance, the practical significance (effect size) of presenting the data with either of our sonification methods is small for  $Q$  (Cohen’s  $|D| \leq 0.19$  for all three comparison pairs, yielding a statistical power of  $(1 - \beta) \leq 0.07$ ). Consequently, a larger sample size would be necessary to detect statistically and practically relevant differences. For  $H$  the effect size is small between S1 and S2 ( $|D| = 0.26, (1 - \beta) = 0.09$ ), medium between S1 and S3 ( $|D| = 0.75, (1 - \beta) = 0.39$ ) and large between S2 and S3 ( $|D| = 1.39, (1 - \beta) = 0.81$ ). These results are inconclusive. On the one hand a larger sample size or a different type of experiment may be necessary to detect and quantify the effect of the sonification design on the participants’ performances. On the other hand this small sample size is already sufficient to detect that S3 is somewhat inferior to S1 and S2 in terms of  $H$ .

We find a statistically and practically significant difference between  $Q$  of our three sonification designs and the rudimentary visualization with high statistical power ( $p = 0.011, |D| = 1.41, (1 - \beta) = 0.87$  for S1,  $p = 0.005, |D| = 1.61, (1 - \beta) = 0.79$  for S2,  $p = 0.008, |D| = 1.52, (1 - \beta) = 0.91$  for S3). Here, the small sample size is already sufficient to allow the conclusion that participants perform significantly using either of the sonification designs compared to rudimentary visualization. Likewise, we find a statistically and practically difference for  $H$  between S1 the rudimentary visualization ( $p = 0.035, |D| = 1.46, (1 - \beta) = 0.89$ ) and between S2 and the rudimentary visualization ( $p = 0.009, |D| = 1.97, (1 - \beta) = 0.99$ ) with high statistical power. Between S3 the rudimentary visualization we only find a statistically non-insignificant, but large effect size and a high power level ( $p = 0.215, |D| = 1.88, (1 - \beta) = 0.98$ ).

Consequently, performance differences between sonification designs and rudimentary visualization are statistically and practically significant and the small sample size is sufficient to produce a high power level.

Comparison between prototypical sonification and our three sonification methods looks very similar. We find statistically and practically significant differences between  $Q$  of all our sonification designs and the prototypical sonification ( $p = 0.045$ ,  $|D| = 1.75$ ,  $(1 - \beta) = 0.97$  for S1,  $p = 0.022$ ,  $|D| = 2.17$ ,  $(1 - \beta) > 0.99$  for S2 and  $p = 0.036$ ,  $|D| = 1.99$ ,  $(1 - \beta) = 0.99$  for S3). Likewise, we find statistically and practically significant differences between  $H$  of all our sonification designs and the prototypical sonification ( $p = 0.108$ ,  $|D| = 1.13$ ,  $(1 - \beta) = 0.70$  for S1,  $p = 0.03$ ,  $|D| = 1.59$ ,  $(1 - \beta) > 0.94$  for S2 and  $p = 0.036$ ,  $|D| = 1.28$ ,  $(1 - \beta) = 0.80$  for S3). Consequently, performance differences between our sonification designs and prototypical sonification are statistically and practically significant and the small sample size is sufficient to produce a high power level.

The rudimentary visualization results do not differ significantly from the prototypical sonification results ( $p = 0.976$  for  $Q$  and  $p = 0.986$  for  $H$ ). The effect size is small for  $Q$  ( $|D| = 0.23$ ,  $(1 - \beta) = 0.08$ ) and medium for  $H$  ( $|D| = 0.47$ ,  $(1 - \beta) = 0.18$ ), and the statistical power is low. A larger sample size would be needed to detect and quantify reliably whether differences between rudimentary visualization and prototypical sonification do exist.

From these results we can conclude that:

- (1) Participants performed significantly better using our sonification designs compared to the rudimentary visualization and the prototypical sonification in terms of  $Q$  and, partly,  $H$
- (2) Participants performed better than chance using either of the data presentation methods
- (3) The effect size of using either of the new sonification designs vs. rudimentary visualization or prototypical sonification is large and the statistical power is high
- (4) The performance when using either rudimentary visualization or prototypical sonification is not significantly different
- (5) Consequently: Rudimentary visualization and prototypical sonification are suitable as a benchmark for  $Q$  and  $H$

#### 4 DISCUSSION

It is not remarkable that our sonification designs tended to enable participants to achieve a higher  $Q$  and  $H$  compared to the prototypical sonification. The limitations of prototypical mappings are certainly important reasons why sonification is not ubiquitous today, despite the audio capabilities of today's technical devices. And these limitations are also the reason why researchers keep conceptualizing and developing new sonification designs. What is more remarkable is that we managed to create a visualization that exhibits many parallels to the presented sonification methods, and that does not produce significantly different results than the prototypical sonification.

We therefore suggest to include the idea of a rudimentary visualization in the search for a suitable benchmark to evaluate your sonification designs, rather than comparing your sonification with

sophisticated, well-established and ubiquitous visualizations. Especially for the study of multi-modal interaction visual benchmarks for the evaluation of auditory designs can be interesting.

Note that the experiment results show that our sonification designs enable users to perform better than the benchmark for the given task. But it does not reveal why. Reasons for a good performance in terms of a  $Q$  and  $H$  could be quick learnability, a low level of sound-induced stress, an intuitive mapping, a high perceptual resolution of the chosen mapping parameters, and/or a high degree of independence between the two mapping dimensions. Further experiments are necessary to identify which of these reasons caused the performance differences. In these experiments, the presented rudimentary visualization and the prototypical sonification could keep serving as benchmarks.

#### 5 CONCLUSION

In this paper we described an experiment in which participants had to assign visualized or sonified data points to fields on a map. We introduced a rudimentary visualization and a prototypical sonification that were supposed to serve as a benchmark.

Experiment results revealed that our sonification designs enabled participants to interpret data points with similar accuracy, which was mostly significantly higher than with rudimentary visualization and prototypical sonification. At the same time, the results based on the usage of rudimentary visualization and prototypical sonification were very similar, and much better than expected by random guessing. This indicates that both may serve as a suitable benchmark to evaluate your sonification design against. A larger sample size and further experiments are necessary to reveal if and to what extent rudimentary visualization and prototypical sonification are equivalent or produce different results.

#### REFERENCES

- [1] David Black, Julian Hettig, Maria Luz, Christian Hansen, Ron Kikinis, and Horst Hahn. 2017. Auditory feedback to support image-guided medical needle placement. *International Journal of Computer Assisted Radiology and Surgery* 12, 9 (2017), 1655–1663. <https://doi.org/10.1007/s11548-017-1537-1>
- [2] Tifanie Bouchara, Brian F. G. Katz, Christian Jacquemin, and Catherine Guastavino. 2010. Audio-Visual Renderings for Multimedia Navigation. In *16th International Conference on Auditory Display (ICAD2010)*, Eoin Brazil (Ed.). Washington, DC, USA, 245–252. <http://hdl.handle.net/1853/58425>
- [3] Gaël Dubus and Roberto Bresin. 2013. A Systematic Review of Mapping Strategies for the Sonification of Physical Quantities. *PLOS ONE* 9, 4 (12 2013), e96018. <https://doi.org/10.1371/journal.pone.0082491>
- [4] Gaël Dubus and Roberto Bresin. 2015. Exploration and evaluation of a system for interactive sonification of elite rowing. *Sports Eng* 18 (2015), 29–41. <https://doi.org/10.1007/s12193-011-0085-1>
- [5] Gustav Theodor Fechner. 1860. *Elemente der Psychophysik*. Breitkopf und Härtel, Leipzig.
- [6] Thomas Hermann. 2002. *Sonification for Exploratory Data Analysis*. Ph. D. Dissertation. Bielefeld University, Bielefeld. <http://www.techfak.uni-bielefeld.de/ags/ami/publications/media/Hermann2002-SFE.pdf>
- [7] Takanori Komatsu and Seiji Yamada. 2016. Can Monaural Auditory Displays Convey Directional Information to Users?. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, A. Papafragou, D. Grodner, D. Mirman, and J. C. Trueswell (Eds.). Cognitive Science Society, Austin, TX, 930–935. <https://mindmodeling.org/cogsci2016/papers/0170/index.html>
- [8] Tapio Lokki and Matti Gröhn. 2005. Navigation with auditory cues in a virtual environment. *IEEE MultiMedia* 12, 2 (April 2005), 80–86. <https://doi.org/10.1109/MMUL.2005.33>
- [9] Frederik Nagel, Fabian-Robert Stöter, Noberito Degara, Stefan Balke, and David Worrall. 2014. Fast and Accurate Guidance – Response Times to Navigational Sounds. In *ICAD*. New York, NY, 5 pages. <http://hdl.handle.net/1853/52058>

- [10] Gaëtan Parseihian, Charles Gondre, Mitsuko Aramaki, Sølvi Ystad, and Richard Kronland-Martinet. 2016. Comparison and Evaluation of Sonification Strategies for Guidance Tasks. *IEEE Trans. Multimedia* 18, 4 (April 2016), 674–686. <https://doi.org/10.1109/TMM.2016.2531978>
- [11] Albrecht Schneider. 2018. Fundamentals. In *Springer Handbook of Systematic Musicology*, Rolf Bader (Ed.). Springer, Berlin, Heidelberg, 559–603. [https://doi.org/10.1007/978-3-662-55004-5\\_30](https://doi.org/10.1007/978-3-662-55004-5_30)
- [12] Roger N. Shepard. 1964. Circularity in Judgments of Relative Pitch. *The Journal of the Acoustical Society of America* 36, 12 (1964), 2346–2353. <https://doi.org/10.1121/1.1919362>
- [13] David Worrall. 2019. *Sonification Design*. Springer, Cham. <https://doi.org/10.1007/978-3-030-01497-1>
- [14] Tim Ziemer, Nuttawut Nuchprayoon, and Holger Schultheis. 2020. Psychoacoustic Sonification as User Interface for Human-Machine Interaction. *International Journal of Informatics Society* 11, 3 (2020), 3–16. <http://www.infso.org/journal/vol11/11-3>
- [15] Tim Ziemer and Holger Schultheis. 2019. Psychoacoustic auditory display for navigation: an auditory assistance system for spatial orientation tasks. *J Multimodal User Interfaces* 13 (2019), 205–218. <https://doi.org/10.1007/s12193-018-0282-2>

## A ONLINE RESOURCES

The visualization and the four sonifications used in this experiment are demonstrated in a YouTube Playlist. They present some coordinates along the axes.