doi:10.20944/preprints202205.0167.v1

PAMPAS: A PsychoAcoustical Method for the Perceptual Analysis of multidimensional Sonification

Tim Ziemer [1,*] **and Holger Schultheis** [1]

[1] *Bremen Spatial Cognition Center, University of Bremen, Bremen , Germany*

Correspondence*:
Tim Ziemer
ziemer@uni-bremen.de

ABSTRACT

The sonification of data to communicate information to a user is a relatively new approach that established itself around the 1990s. To date many researchers design their individual sonification from scratch. There are no standards in sonification design and evaluation. But researchers and practitioners have formulated several requirements and established several methods. There is wide consensus that psychoaocustics could play an important role in the sonification design and evaluation phase. But this requires an adaption of psychoacoustic methods to the signal types and the requirements of sonification. In this method paper we present PAMPAS, a PsychoAcoustical Method for the Perceptual Analysis of multidimensional Sonification. A well-defined and well-established, efficient, reliable and replicable Just Noticeable Difference experiment using the Maximum Likelihood Procedure serves as the basis to achieve linearity of parameter mapping during the sonification design stage and to identify and quantify perceptual effects during the sonification evaluation stage, namely the perceptual resolution, hysteresis effects and perceptual interferences. The experiment results are universal scores from a standardized data space and a standardized procedure. These scores can serve to compare multiple sonification designs of a single researcher, or even between different research groups. The method can supplement other sonification design and evaluation methods from a perceptual viewpoint.

Keywords: sonification evaluation, psychoacoustics, just noticeable difference, difference limen, discrimination threshold, comparison of sonification designs, maximum likelihood procedure, auditory display

1 INTRODUCTION

Sonification is the systematic conversion from data to sound so that aspects of the data become unambiguously understandable for a user Hermann (2008, 2021); Scaletti (2018). Even though antecedents can be found throughout all periods of history, *sonification* as a term and as a dedicated field of research established itself around the early 1990s (Worrall, 2019, chap. 1).

Sometimes sonified data is multidimensional. Examples include an abstract phase space Hermann (2018), real spatial locations Lokki and Gröhn (2005); Ziemer and Schultheis (2019a); Ziemer et al. (2020) and angles Asendorf et al. (2021); Greindl et al. (2020). Sometimes, data is multivariate. Examples include vital functions of patients, like pulse frequency and blood oxygen concentration Fitch and Kramer (1993); Watson and Sanderson (2004); Yeung (1980); Ziemer et al. (2020), geopolitical data, like crime rate and unemployment rate Olivetti Belardinelli et al. (2009), or pH and chrlorine level in water Ziemer et al. (2020). Throughout this paper we will refer to both cases as being *multidimensional*.

In the sonification literature, there is consensus that each dimension needs to

- be linear (Hermann, 2002, p. 39)(Worrall, 2019, p. 42)Ziemer and Schultheis (2020); Ziemer et al. (2020)
- exhibit a high perceptual resolution Ziemer et al. (2020); Ziemer and Schultheis (2019b); Vickers (2006); Yeung (1980); Brewster (2003); Ziemer and Schultheis (2020) (Hermann, 2002, pp. 34 & 40)
- exhibit preferably little hysteresis effects Neuhoff and Kramer (2000); Neuhoff (1998); Feron et al. (2009); Martens (2002)
- exhibit little perceptual interference with other dimensions (Worrall, 2019, p. 42)(Hermann, 2002, p. 34)Ziemer et al. (2020); Anderson and Sanderson (2009); Neuhoff and Kramer (2000); Ziemer and Schultheis (2019b); Barrass and Kramer (1999); Yeung (1980); Ziemer and Schultheis (2020).

Unfortunately, the literature does not provide any guidelines to achieve a multi-dimensional sonification that fulfills these requirements. Worrall (2014) discusses the mapping problem and highlights the need for heuristic testing to approach multi-dimensional sonification. Neuhoff et al. (2002) underline the need for interdisciplinary research for developing and evaluating auditory displays. Ferguson et al. (2006) even suggest the use of psychoacoustic models in the design and evaluation phase of multidimensional sonification. Still, utilizing psychoacoustic methods is not common practice in sonification research. It has been pointed out that the psychoacoustic literature neither deals with the sounds to be expected in sonification design, nor aims at solving the problems that need to be assessed and overcome in sonification design Bonebright and Flowers (2011); Bovermann et al. (2011); Walker and Nees (2011); Ferguson et al. (2006); Smith (1990); Wegner and Karron (1997); Smith et al. (1994); Anderson and Sanderson (2009); Vogt (2011); Barrass (1997).

In this method paper we propose PAMPAS, a **P**sycho**A**coustic **M**ethod for the **P**erceptual **A**nalysis of multidimensional **S**onification. PAMPAS serves as an aid to achieve linearity during the sonification design stage and as a tool to assess the perceptual resolution, hysteresis effects and perceptual interference during the sonification evaluation stage. This is achieved by adapting a psychoacoustic experiment to the signals and requirements of multi-dimensional sonification.

The remainder of this paper is structured as follows: We start with the state-of-the-art in sonification design and evaluation in Section 2. The section summarizes the common practice and the discussions about the potentials and limitations of psychoacoustic methods for sonification research. Section 3 lists the materials and equipment needed to carry out the proposed psychoacoustic experiment. Section 4 describes the proposed psychoacoustic method that is supposed to overcome current limitations of psychoacoustic procedures by being optimized for the signal types and the requirements to be expected in multi-dimensional sonification. Section 5 demonstrates how to extract and report the expected experiment results. Section 6 concludes the work, discusses its strengths and weaknesses and gives an outlook on further psychoacoustic methods that could be adapted to become utilizable for sonification research.

2 STATE OF THE ART

In the field of sonification research many works highlight that there is a need for, but a lack of, comprehensive sonification design guidelines Ibrahim et al. (2011); Supper (2012); Neese and Walker

(2009). In contrast to that, many sonification evaluation methods, especially listening tests Seiça et al. (2020), have proven their value in several studies.

2.1 Conventional Sonification Design and Evaluation

The chapter on Auditory Display Evaluation in The Sonification Handbook suggest that experts should use their own introspection and intuition, especially in the early stage of sonification design (Bonebright and Flowers, 2011, p. 112), and Vogt (2011) agrees to the necessity for sonification researchers to evaluate their initial sonification designs subjectively. After this early sonification design stage, “a variety of methods including both laboratory components and ecologically valid testing” (Bonebright and Flowers, 2011, p. 111) is recommended.

Typically, sonification evaluations with potential end-users are presented in the sonification literature Seiça et al. (2020), rather than initial self-tests Ibrahim et al. (2011). To evaluate their sonification design, many researchers let users solve a task with auditory, visual, and audiovisual guidance Lokki and Gröhn (2005); Black et al. (2017); Ziemer and Schultheis (2019a), others compare multiple sonification designs for a specific task Parseihian et al. (2016); Komatsu and Yamada (2016); Albrecht et al. (2016); Walker and Lindsay (2006); Rönnberg (2019). In navigation studies researchers evaluate for example completion time Lokki and Gröhn (2005); Parseihian et al. (2016); Hansen et al. (2013); Ziemer and Schultheis (2019a); Albrecht et al. (2016); Walker and Lindsay (2006), precision Parseihian et al. (2016); Hansen et al. (2013); Ziemer and Schultheis (2019a), accuracy Ziemer and Schultheis (2019a), turn arounds Parseihian et al. (2016); Ziemer and Schultheis (2019a); Walker and Lindsay (2006), interruptions Parseihian et al. (2016); Ziemer and Schultheis (2019a), trajectory lengths Ziemer and Schultheis (2019a); Walker and Lindsay (2006), trajectory entropy Ziemer and Schultheis (2019a), qualitative inspection of trajectories Lokki and Gröhn (2005); Ziemer and Schultheis (2019a); Albrecht et al. (2016), and training effects Nagel et al. (2014); Walker and Lindsay (2006); Ziemer and Schultheis (2019a). Sometimes, navigation is evaluated in a game-like scenario to in the hope for a high motivation of long-term usage Degara et al. (2013); Ziemer and Schultheis (2021); Pires et al. (2013); and (2019).

In addition to analyzing task performances, researchers often determine the subjective task load using the NASA-Task Load Index (TLX) Ziemer and Schultheis (2019a); Khan and Jeon (2018) or the Raw NASA-TLX Black et al. (2017); Ziemer and Schultheis (2019a); and (2019) and use subjective questionnaires concerning, e.g., aesthetics Kuppanda et al. (2015); Vogt (2011); Neumann et al. (2013); Vickers (2006), annoyance Ziemer et al. (2018); Vickers (2006); Brewster (2003), clarity Vogt (2011); Pauletto and Hunt (2007), comprehensibility Yang and Hunt (2013); Neumann et al. (2013); Vickers (2006), distracting Neumann et al. (2013); Vickers (2006), informativeness Ziemer et al. (2018); Neumann et al. (2013); Hermann et al. (2015); Khan and Jeon (2018), intuitiveness Kuppanda et al. (2015); Vogt (2011); Khan and Jeon (2018); Pauletto and Hunt (2007), learnability/learning effort Neumann et al. (2013); Vogt (2011), obtrusiveness Neumann et al. (2013), pleasantness Kuppanda et al. (2015); Neumann et al. (2013); Hermann et al. (2015); Pauletto and Hunt (2007), preference Yang and Hunt (2013), and/or utility/usefulness Khan and Jeon (2018); Neumann et al. (2013).

In summary, sonification evaluations tend to include task-solving, NASA-TLX measurement, and subjective questionnaires. In addition to these sonification design and evaluation methods, the need for a psychoacoustically well-informed sonification design has been raised Bovermann et al. (2011); Walker and Nees (2011); Brewster (2003) and common practices from the field of psychoacoustics, such as identification, attribute rating, and discrimination tasks (Zwicker and Fastl, 1999, chap. 1), have been suggested to evaluate sonification designs Bonebright and Flowers (2011); Walker and Nees (2011).

2.2 Psychoacoustic Sonification Design and Evaluation

Many studies highlight the difficulty of psychocoustic sound synthesis Ferguson et al. (2006); Smith (1990); Wegner and Karron (1997); Smith et al. (1994), because psychoacoustic models like Leman (2000); Zwicker and Fastl (1999); Sottek (2016) solve the forward problem, whereas psychoacoustic sound synthesis has to solve the inverse problem. Here, the forward problem is: An acoustic input is given and you transform it to predict the perceptual outcome. The inverse problem is: The desired perceptual outcome is given, and you need to find an acoustical input that will produce this perception. Given the complex and nonlinear transformation that our auditory system performs on the acoustic input, there is neither an analytical nor a comprehensive numerical solution to the inverse problem. The only exception are very primitive signals. For example the Mel scale (Zwicker and Fastl, 1999, chap. 5)Schneider (2018a) and the phon scale (Zwicker and Fastl, 1999, chap. 8)Schneider (2018a) can serve as lookup-tables to produce the desired pitch and loudness by manipulating frequency and amplitude of pure tones within certain limits.

The psychoacoustic literature has been argued to be of limited relevance for sonification, because psychoacousticians tend to use test stimuli that differ a lot from sonification Anderson and Sanderson (2009); Vogt (2011); Smith et al. (1994), like pure tones, Gaussian-shaped tone burst, band-pass noise and DC-pulses (Zwicker and Fastl, 1999, chap. 1). Consequently, the results of psychoacoustic experiments, like Just Noticeabe Difference (JND) experiments, are not necessarily valid for a sonification. Even though the number of studies that applied psychoacoustic methods on sonification—like magnitude estimation Walker and Kramer (2005); Walker (2002) and similarity ratings Bonebright (2001); Fernstrom et al. (2003)—is low, there is the belief that "psychoacoustic measurements and theories can assist in the design of an auditory display" (Barrass, 1997, pp. 17–18).

For multi-dimensional sonification there is consensus that researchers need to ensure linearity Ziemer et al. (2020) (Barrass, 1997, p. 115) (Hermann, 2002, p. 141), and evaluate the perceptual resolution Ziemer et al. (2020); Vickers (2006); Ziemer and Schultheis (2019b) (Hermann, 2002, p. 34) (Barrass, 1997, chap. 7.7.5), hysteresis effects Neuhoff and Kramer (2000); Neuhoff (1998); Feron et al. (2009); Martens (2002), and perceptual interferences Ziemer and Schultheis (2019c); Brewster (2003); Ziemer et al. (2018, 2020); Ziemer and Schultheis (2019b); Neuhoff (2019); Worrall (2014); Kramer (1994); Neuhoff and Kramer (2000); Ziemer et al. (2020); Ziemer and Schultheis (2019b); Neuhoff et al. (2002); Anderson and Sanderson (2009); Barrass and Kramer (1999); Ferguson et al. (2006); Barrass and Kramer (1999) (Worrall, 2019, chap. 2.2.2.2)(Hermann, 2002, p. 34)(Barrass, 1997, p. 16) between dimension. These are briefly described in the following.

2.2.1 Linearity

The necessity of linearity in sonification has been expressed, e.g., in Ziemer et al. (2020) (Barrass, 1997, p. 115) (Hermann, 2002, p. 141). In a linear system the output is directly proportional to the input, so that the doubling of an input value doubles the output value. In psychophysics, a logarithmic transform of the physical input sometimes produces a fairly linear sensational output (Fechner, 1860, pp. 134ff)Schneider (2018a). For exampling a constant *frequency ratio* tends to sound like a fairly constant *pitch interval*. According to the psychoacoustic literature (Fechner, 1860, p. 60)Schneider (2018a), a linear sensory scale is achieved when the JND is constant throughout the scale.

2.2.2 Perceptual Resolution

The necessity to achieve a sufficient resolution, and to quantify the perceptual resolution of sonification designs has been expressed, e.g., in Ziemer et al. (2020); Vickers (2006); Ziemer and Schultheis (2019b) (Hermann, 2002, p. 34) (Barrass, 1997, chap. 7.7.5).

The JND is an adequate measure of perceptual resolution, which is acquired in psychoacoustic experiments. Here, all but one physical parameter tend to be kept constant, so the JND of the remaining parameter can be measured. For example the JND in frequency of pure tones evolves from 1% at low frequencies over 0.3% at midrange frequencies to 1% at high frequencies (Scheminzky, 1943, p. 136)(Ziemer, 2020, chap. 4). The JND in amplitude of a pure tone is 0.3 to 1.4 dB, depending on reference amplitude and frequency (Ziemer, 2020, pp. 74f).

In sonification, audio signals tend to be more complex and unique compared to test signals in conventional psychoacoustic experiments. Consequently, the JNDs reported in the psychoacoustic literature are no reliable indicators of perceptual resolution of sonification dimensions.

2.2.3 Hysteresis effects

Hysteresis effects are well-known in the field of psychoacoustics and even exist for fundamental aspects of sound perception, such as pitch Stevens and Volkmann (1940); Greenwood (1997); Chambers and Pressnitzer (2014) and loudness Canévet and Scharf (1990). Hysteresis is the effect that the distance from data point $\mathbf{a}$ to $\mathbf{b}$ may be perceived different than the distance from $\mathbf{b}$ to $\mathbf{a}$. This means the difference not only depends on the interval but also on where you're coming from.

Sonification researchers mention the existence of perceptual hysteresis effects, too Neuhoff and Kramer (2000); Neuhoff (1998); Feron et al. (2009); Martens (2002). But it is typically neither quantified nor qualitatively discussed in the sonification literature.

2.2.4 Perceptual Interference

Many works on sonification address the issue of orthogonality in multidimensional sonification Ziemer and Schultheis (2019c); Brewster (2003); Ziemer et al. (2018, 2020); Ziemer and Schultheis (2019b); Neuhoff (2019); Worrall (2014); Kramer (1994); Neuhoff and Kramer (2000); Ziemer et al. (2020); Ziemer and Schultheis (2019b); Neuhoff et al. (2002); Anderson and Sanderson (2009); Barrass and Kramer (1999); Ferguson et al. (2006); Barrass and Kramer (1999) (Worrall, 2019, chap. 2.2.2.2)(Hermann, 2002, p. 34)(Barrass, 1997, p. 16), i.e., interference or interactions between dimensions. This means that changes along one data dimension sound like changes along two or more data dimensions, or that the absolute magnitude of data dimension one affects the resolution of data dimension two. Perceptual interference can "obscure data relations and confuse the listener" (Worrall, 2019, chap. 2.2.2.2).

In the field of psychoacoustics it is accepted that all physical aspects of sound can affects practically all aspects of sound perception. For example even though amplitude and frequency of a pure tone are physically orthogonal, they both interfere perceptually, as both can affect the sensation of, e.g., loudness and pitch (Zwicker and Fastl, 1999, chaps. 5.1.2 and 8.1) Schneider (2018b). In the psychoacoustic literature, fundamental aspects of sound sensation like loudness, roughness, sharpness and tonalness are considered largely independent from one another Aures (1985) (Zwicker and Fastl, 1999, chap. 9) albeit all of them can be influenced by very many physical sound parameters to some extent. But as mentioned above, the psychoacoustic literature concentrates on the forward problem and has not yet offered a solution to the inverse problem, which would be needed for a psychoacoustic sonification design.

3 MATERIALS AND EQUIPMENT

To carry out the PAMPAS on your sonification design, you need a computer with a internal or external sound card, high-quality D/A-converter, amplifier and headphones. Ideally, you calibrate your system using an artificial ear or an artificial head that informs you about the sound pressure level to be expected at the participants ear drums, which should lie around 70 dB.

For remote testing, an online survey platform is needed. This platform either needs to be able to play uncompressed sound files, like wav or aiff, or to render audio on the fly. More importantly, it has to be adaptive, i.e., it must be able to play a specific audio signal that depends on the user's previous actions.[1]

Should your sonification design require any spatialization techniques — such as vbap, ambisonics or wave field synthesis Ziemer (2020) — you need a respective loudspeaker setup and document the technical details on the setup.

4 METHODS

In psychophysical terms, PAMPAS is based on a Just Noticeable Difference (JND) experiment using the Maximum Likelihood Procedure (MLP) in a 2 Alternative Forced Choice (2AFC) task. These kinds of experiments are described in detail, e.g., in Grassi and Soranzo (2009); Madigan and Williams (1987); Green (1990, 1993b).

This section starts with the aim of the presented experiment, followed by the necessary steps to prepare your data and signals, and to conduct the experiment.

4.1 Aim

The aim of the proposed experiment is two-fold: 1. During the sonification design stage a single researcher or a small group of researchers and sound designers can carry out the "light" experiment to achieve a perceptually *linear* mapping. 2. After implementing a prototype, the multi-dimensional sonification can be evaluated in experiments with participants from the target audience to quantify

1. the *perceptual resolution* of,
2. *hysteresis effects* within, and
3. *perceptual interferences* between

each dimension and polarity. The experiment results are quantitative measures that help to assess aspects of an individual sonification design and to compare multiple sonifications.

4.2 Data Space

We start with an abstract, normalized data space that has two Cartesian dimensions x and y, as illustrated in Fig. 1. We refer to this as the *mapping space* that contains the *mapping data*. Each dimension has two *polarities*, a negative and a positive one, so that the dimensions range from -1 to 1. This allows us to sonify data in interval scale. We can divide the space into 4 quadrants and name them roman I to IV. Should the sonification design be conceptualized for data in ratio scale, i.e., without a negative polarity, the data

---

[1] Amongst others, MATLAB Online / MATLAB Web App Server fulfills these demands. It is one of the platforms suggested by the Task Force on Remote Testing by the Technical Committee on Psychological and Physiological Acoustics of the Acoustical Society of America: `https://www.spatialhearing.org/remotetesting/Main/HomePage` and an MLP toolbox for MATLAB is freely availabe under `https://dpg.unipd.it/en/mlp/mlp-toolbox`.

space only contains quadrant I. Naturally, a dataspace can also include combinations of uni- and bipolar dimensions, i.e., quadrants I and II or quadrants I and IV.

In the center of each quadrant we have one reference coordinate, referred to as the *standard* ($S_{\text{I to IV}} = (\pm 0.5, \pm 0.5)$). Sounds evaluated against the standard are called the *variables*. From every standard, we can either move along the positive or negative x-direction ($\pm \Delta x$), the positive or negative y-direction ($\pm \Delta y$) or the positive diagonal ($+\Delta x / + \Delta y$). Only these 5 motions will be tested during the experiment to keep it short.

The normalized data space is necessary in order to compare a number of evaluated sonifications, independent of their intended use case.

4.3 Data Normalization

The abstract data space does not have to equal the data space for the intended sonification use case. If you designed your sonification to deal with values that lie well outside the range of the data space, you need to carry out a normalization.

For example you may want to sonify if and how much the temperature t and the chlorine level c of the water in your hot tub deviate from your target temperature $t_{\text{tar}} = 39°\text{C}$ and chlorine level $c_{\text{tar}} 3$ ppm. Acceptable temperature ranges from $t_{\text{min}} = 35°\text{C}$ to $t_{\text{max}} 42°\text{C}$ and chlorine levels from $c_{\text{min}} = 1$ ppm to $c_{\text{max}} 4$ ppm. In this case you normalize your sensor data in relation to your target values to conform with the data space via

$$x = \begin{cases} 1, \text{ if } t \geq t_{\text{max}} \\ \frac{t - t_{\text{tar}}}{t_{\text{max}} - t_{\text{tar}}}, \text{ if } t_{\text{tar}} < t < t_{\text{max}} \\ \frac{t - t_{\text{tar}}}{t_{\text{tar}} - t_{\text{min}}}, \text{ if } t_{\text{min}} < t \leq t_{\text{tar}} \\ -1, \text{ if } t \leq t_{\text{min}} \end{cases} \tag{1}$$

and

$$y = \begin{cases} 1, \text{ if } c \geq c_{\text{max}} \\ \frac{c - c_{\text{tar}}}{c_{\text{max}} - c_{\text{tar}}}, \text{ if } c_{\text{tar}} < c < c_{\text{max}} \\ \frac{c - c_{\text{tar}}}{c_{\text{tar}} - c_{\text{min}}}, \text{ if } c_{\text{min}} < c \leq c_{\text{tar}} \\ -1, \text{ if } c \leq c_{\text{min}} \end{cases} \tag{2}$$

Equations 1 and 2 are transforms from the sensed or defined *input data* to the abstract, normalized *mapping data*. The transform is also referred to as *data normalization* or *piecewise linear transfer function* (Hermann, 2002, p. 38).

Naturally, the data normalization can be preceded by additional transforms. For example it may be meaningful to transform from the measured concentration of hydrogen ions in a liquid to the logarithmic pH value.

4.4 General Procedure

The general procedure of a JND experiment using the 2AFC approach is simple. First, the conductor explains the participant that a new type of informative sound is being tested. It is wise to present the sonification of discrete coordinates in ascending and descending order and visualize them on a graph similar

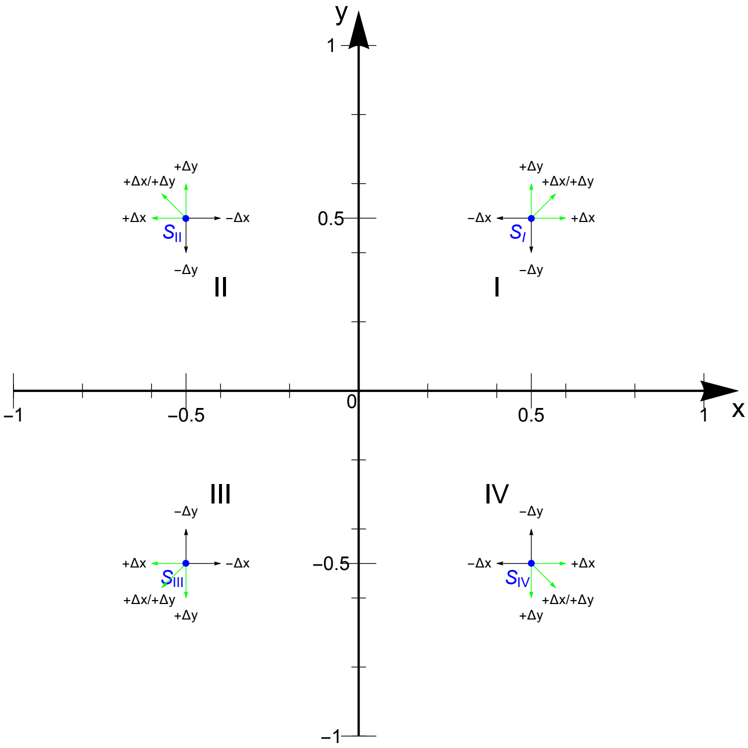

Figure 1. The mapping space. It has two linear dimensions that are normalized from -1 to 1. The MLP is carried out in 5 directions (small arrows) at 4 locations (S_{I} to S_{IV}) to quantify the perceptual resolution, hysteresis effects, and perceptual interferences. The standards lie in the center of each quadrant.

to Figure 1, and, ideally even let the participants explore the sonification interactively themselves, using a computer mouse or any other familiar human interface device. This way the participants can familiarize with the sound, which characteristics it has, and how small or large the intensity or magnitude of each characteristics can be. To test the new type of sound, sound pairs will be presented to the participant, who has to judge which of the two has the larger intensity, magnitude, or significance of a certain characteristic, i.e., which one lies further away from the center of the coordinate origin in the respective example. They have to do this over and over, for different regions within the presented graph.

4.5 Signal Preparation

In this section the experiment will be explained for quadrant I. The same procedure has to be carried out for quadrants II to IV, respectively.

First, you produce the sound of the standard, i.e., you sonify coordinate $(0.5, 0.5)$. Next, you produce the variable test sounds. For the positive x-direction, these are $(0.5 + \alpha, 0.5)$, where α goes from 0.001 to 0.1 in steps of 0.001, yielding $J = 100$ test sounds. Likewise, you produce 100 test sounds for the negative x-direction $(0.5 - \alpha, 0.5)$, for the positive y-direction $(0.5, 0.5 + \alpha)$, the negative y-direction $(0.5, 0.5 - \alpha)$ and the diagonal $(0.5 + \alpha, 0.5 + \alpha)$. The sounds should last for 3 to 5 seconds, followed by a pause of equal length.

Note that discrete test sounds of finite duration may be different from the sound of the sonification in the intended use case, where it may be continuous. That is why you need to ensure that the test sounds do not provide audible cues that would not appear during the usual sonification usage. For example you may need to ramp the signal to avoid audible clicks at the onset and offset of the sound. Imagine the x value was mapped to the frequency of a Low Frequency Oscillator (LFO) that modulated the frequency of a carrier, creating a vibrato effect. Here, you should ensure that the sound of the standard and the variable start at the same phase of the LFO cycle. This way the listener's judgments will not be biased by the phase at the onset of the two sounds. Furthermore, you should choose a long ramp at the offset. This gradual fade-out will ensure that listeners do not compare the two sounds based on the phase of the LFO cycle during the note offset.

4.6 Psychometric Functions

A psychometric function describes how likely a participant can distinguish two stimuli. This function is sampled for each individual during a 2AFC experiment and can be expressed as

$$\Psi(x;\alpha,\beta,\gamma) = \gamma + (1-\gamma) f(x;\alpha,\beta)\ . \tag{3}$$

Here, the chance to guess correctly lies at $\gamma = 50\%$ and the function should converge to this threshold towards the lower end, the *floor*. Ideally, the *ceiling* should be 100%, even though an *attentional lapse* could reduce the ceiling a bit.

The function $f(x;\alpha,\beta)$ approximates the sampled psychometric function by a sigmoid function, like the logistic function

$$f_{\text{logistic}}(x;\alpha,\beta) = \frac{1}{1+e^{\beta(\alpha-x)}}\ , \tag{4}$$

where α is the midpoint on the x-axis and and β manipulates the slope of the curve.

4.7 Maximum Likelihood Procedure

Before starting the experiment, several psychometric functions are formulated, referred to as *hypotheses* H_j. These have the same slope β and chance level γ, but different midpoints α that spread over the complete range in which the JND is assumed to lie. An example is given in Fig. 2.

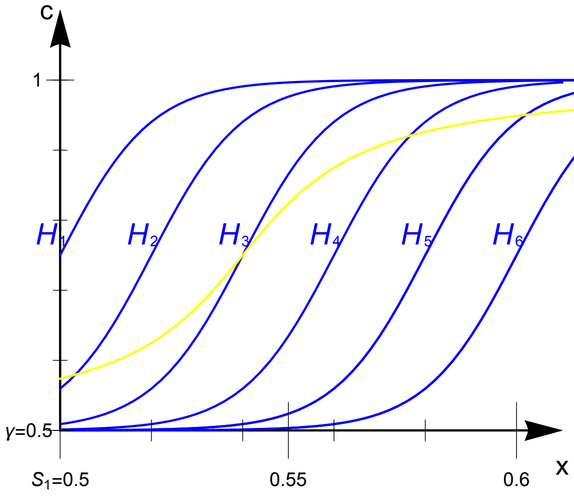

Figure 2. Six exemplary hypotheses (blue) and the actual psychometric function (yellow). Even though the slope is different, H_3 approximates the JND well.

doi:10.20944/preprints202205.0167.v1

Table 1. Proposed values for a standardized MLP to evaluate two-dimensional sonification designs.

standard	γ	$f(x;\alpha,\beta)$	β	α	p_{target}	J hypotheses	trials
$(\pm 0.5, \pm 0.5)$	50%	$\frac{1}{1+e^{\beta(\alpha-x)}}$	100	$0.001 \leq \alpha \leq 0.1$	80.9%	100	12

After every trial n, the likelihood of each hypothesis to equal the participant's psychometric function can be calculated as

$$L(H_j) = \sum_{i=1}^{n} C \lg H(x_i) + W \lg (1 - H(x_i)) \tag{5}$$

where x is the stimulus magnitude and the index i addresses all previous trials up to the latest trial. Here, the exponents C equals 1, if the participants gives the correct answer, otherwise it is 0. Respectively, W equals 1 if the answer is wrong and 0 if the answer is correct. The hypothesis with the highest likelihood gives the best approximation of a subject's psychometric curve with the answered given from trial 1 to the nth trial. The calculated likelihood is recalculated every trial and becomes more accurate, so the last trial of the whole experiment returns the best estimate.

4.8 Stimulus Selection Method

It is certainly motivating to start with a clearly audible difference, like the sound of coordinate $(0.57, 0.5)$. After each trial, the subject's most likely JND is the inverse function of his or her psychometric function at its p_{target}, i.e.,

$$\Psi^{-1}(p_{\text{target}}) = \alpha_j - \frac{1}{\beta} \ln \left(\frac{1-\gamma}{p_{\text{target}} - \gamma} - 1 \right) . \tag{6}$$

This calculated p_{target} will be the stimulus for the next trial. This way we ensure that most trials in the experiment are near the JND. In a 2AFC task p_{target} equals 80.9%, which minimizes the variability in the estimate of the threshold Green (1993a).

4.9 Conventions

Maximum likelihood procedures allow for many decisions to adapt it to the given task and the specific objective of an experiment. For sonification evaluation the main objective is to quantify

1. the perceptual resolution (JND) of,
2. hysteresis effects (h) of and
3. perceptual interferences (δ) between

x and y in quadrants $S_{\text{I to IV}}$.

Therefore, we suggest to use the values indicated in Table 1 that allow for a quick, reliable and comparable evaluation of two-dimensional sonification designs.

However, a premise to evaluate a multi-dimensional sonification with potential end-users utilizing the proposed MLP is that the sonification dimensions sound fairly linear. To assist in achieving linearity, the procedure should be carried out by the sonification designers and some colleagues with standards that equally sample the axes of the normalized data space, i.e., the tick marks in Figure 1. This "light" experiment is an aid for sonification design.

doi:10.20944/preprints202205.0167.v1

Table 2. Table to report the perceptual resolution of each dimension in each quadrant.

	I	II	III	IV
x	$\text{JND}_{S_\text{I}}(+\Delta x)$	$\text{JND}_{S_\text{II}}(+\Delta x)$	$\text{JND}_{S_\text{III}}(+\Delta x)$	$\text{JND}_{S_\text{IV}}(+\Delta x)$
y	$\text{JND}_{S_\text{I}}(+\Delta y)$	$\text{JND}_{S_\text{II}}(+\Delta y)$	$\text{JND}_{S_\text{III}}(+\Delta y)$	$\text{JND}_{S_\text{IV}}(+\Delta y)$

4.10 Demonstrations to the participants

Before you start the actual experiment, you should present some audio examples to the participants. We suggest to sonify the coordinate $(0.6, 0.6)$ to them as a reference, $(0.601, 0.6)$ as a “probably inaudible difference”, $(0.7, 0.6)$ as a “probably clearly audible difference” and then a some coordinate near your personal JND, just to give them the feel for the level of nuances they should try to achieve.

5 ANTICIPATED RESULTS

At each standard $S_{\text{I to IV}}$ the MLP yields 5 just noticeable differences, namely

1. along the x-direction, towards $x = 0$ (i.e., $\text{JND}_S(-\Delta x)$) and
2. away from $x = 0$ (i.e., $\text{JND}_S(+\Delta x)$),
3. along the y-direction, towards $y = 0$ (i.e., $\text{JND}_S(-\Delta y)$) and
4. away from $y = 0$ (i.e., $\text{JND}_S(+\Delta y)$),
5. and along its diagonal (i.e., $\text{JND}_S(+\Delta x, +\Delta y)$.

These JNDs inform about the perceptual resolution, the presence, direction and intensity of hysteresis effects, and the degree of perceptual interference as explained below.

5.1 Perceptual Resolution

In each quadrant the perceptual resolution is described by two JNDs, namely the JND when increasing the absolute value of either x or y, i.e., $\text{JND}_S(+\Delta x)$ and $\text{JND}_S(+\Delta y)$. These directions are indicated as the orthogonal, green arrows in Figure 1.

For example the $\text{JND}_{S_\text{I}}(+\Delta x)$ reveals the perceptual resolution of the x-dimension in quadrant I. Ideally, these JNDs are very small, indicating a high perceptual resolution of the respective dimension. In the proposed method the smallest possible value is 0.001, which would indicate that $1/0.001 = 1000$ steps can be distinguished by listeners in quadrant I, given that the parameter mapping along the respective dimension is truly linear from $x = 0$ to $x = 1$. The highest possible JND is 0.1, which would indicate that $1/0.1 = 10$ steps can be distinguished by listeners. This score is very low. Such a sonification is not suitable for the presentation of a continuous axis in interval or ratio scale, but rather for ordinal or nominal data.

The perceptual resolution should be reported as a table, as demonstrated in Table 2.

Preceding the sonification evaluation experiment, sonification designers should ensure linearity of the axis by measuring $\text{JND}(+\Delta x)$ at each tick mark along the x-axis and $\text{JND}(+\Delta y)$ at each tick mark along the y-axis in a “light” self-experiment. Here, the JND at every tick mark of each axis polarity should be equal. If this is not the case, the parameter mapping function needs to be adjusted.

Table 3. Table to present the hysteresis effects of each dimension in each quadrant

	I	II	III	IV
x	$h_{S_{\mathrm{I}}}(x)$	$h_{S_{\mathrm{II}}}(x)$	$h_{S_{\mathrm{III}}}(x)$	$h_{S_{\mathrm{IV}}}(x)$
y	$h_{S_{\mathrm{I}}}(y)$	$h_{S_{\mathrm{II}}}(y)$	$h_{S_{\mathrm{III}}}(y)$	$h_{S_{\mathrm{IV}}}(y)$

5.2 Hysteresis effects

In each quadrant the ratio between the JND along the positive and along the negative direction of a dimension informs us about hysteresis effects. Ideally, $\mathrm{JND}_S(-\Delta x)$ equals $\mathrm{JND}_S(+\Delta x)$, meaning that no hysteresis effects exists along x in the respective quadrant. If either JND is greater, it implies that larger changes along that direction are necessary in order to be audible. As a fairly linear mapping is a prerequisite for the listening experiments, the ratio between the two JNDs indicates the strength of and direction of a hysteresis effect, calculated as

$$h_S(x) = 0.5 \lg \frac{\mathrm{JND}_S(-\Delta x)}{\mathrm{JND}_S(+\Delta x)} \;. \tag{7}$$

The fraction in Equation 7 yields a positive value that lies between $0.001/0.1 = 10^{-2}$ and $0.1/0.001 = 10^{2}$, so the score $h_S(x)$ is normalized to values between -1 and 1. Here, positive values mean that when approaching the coordinate origin, larger steps are necessary in order to be audible, compared to the situation in which you move away from the coordinate origin. In other words: positive values indicate that the perceptual resolution in the direction of the coordinate origin is higher compared to the other direction along the respective dimension. The larger the absolute value of $h_S(x)$ the stronger this hysteresis effect. Naturally, a respective $h_S(y)$ quantifies the hysteresis effects along the y-dimension.

The hysteresis scores should be reported as a table for each quadrant and direction, as demonstrated in Table 3.

Whether hysteresis effects are acceptable or problematic depends on the intended use case.

5.3 Perceptual interference

The diagonal JND, i.e., $\mathrm{JND}_S(+\Delta x, +\Delta y)$, is an indicator for perceptual interference, especially in relation to the JNDs of the single dimensions $\mathrm{JND}_S(+\Delta x)$ and $\mathrm{JND}_S(\Delta y)$. These are indicated as green arrows in Figure 1.

First, we define the minimum $\mathrm{JND}_{S_{\mathrm{min}}}$ as the smaller value of $\mathrm{JND}_S(+\Delta x)$ and $\mathrm{JND}_S(+\Delta y)$, i.e.,

$$\mathrm{JND}_{S_{\mathrm{min}}} = \min\left(\mathrm{JND}_S(+\Delta x), \mathrm{JND}_S(+\Delta y)\right) \tag{8}$$

and the maximum $\mathrm{JND}_{S_{\mathrm{max}}}$ as the larger one, i.e.,

$$\mathrm{JND}_{S_{\mathrm{max}}} = \max\left(\mathrm{JND}_S(+\Delta x), \mathrm{JND}_S(+\Delta y)\right) \;. \tag{9}$$

Their ratio is our threshold T defined as

$$T_S = 0.5 \lg \frac{\mathrm{JND}_{S_{\mathrm{max}}}}{\mathrm{JND}_{S_{\mathrm{min}}}} \;. \tag{10}$$

Table 4. Table to report the perceptual interference in each quadrant.

	δ	Q
S_{I}	$\delta_{S_{\mathrm{I}}}$	$Q(\delta_{S_{\mathrm{I}}})$
S_{II}	$\delta_{S_{\mathrm{II}}}$	$Q(\delta_{S_{\mathrm{II}}})$
S_{III}	$\delta_{S_{\mathrm{III}}}$	$Q(\delta_{S_{\mathrm{III}}})$
S_{IV}	$\delta_{S_{\mathrm{IV}}}$	$Q(\delta_{S_{\mathrm{IV}}})$

Equation 10 is similar to our hysteresis definition, Equation 7. But as min $\leq$ max, T_S can only take values from 0 to 1.

The degree of perceptual interference δ is defined as

$$\delta_S = 0.5\lg\frac{\mathrm{JND}_S(+\Delta x, +\Delta y)}{\mathrm{JND}_{S_{\mathrm{min}}}}\,. \tag{11}$$

Here, the score δ_S could theoretically take values from -1 to 1. Qualitatively, we can distinguish four cases of perceptual interference $Q(\delta)$, namely

$$Q(\delta) = \begin{cases} \text{positive interference, if } \delta_S < 0 \\ \text{no interference, if } \delta_S = 0 \\ \text{usual interference, if } 0 < \delta_S \leq T_S \\ \text{negative interference, if } \delta_S > T_S \end{cases}\,. \tag{12}$$

Positive interference is a rare case in which changes along one dimension improve the user's precision in detecting changes along the second dimension. No interference implies that the two dimensions are orthogonal, which is also very rare not only in audition, but in psychology in general. The usual interference implies that the JND along the diagonal lies somewhere between the JND of the single dimensions. If the JND of the diagonal is even larger than the larger of the two single-dimension JNDs, we refer to that as negative interference.

The quantitative and qualitative interference of all quadrants should be reported in a table, as shown in Table 4.

6 DISCUSSION

In this paper we gave an overview about sonification design and evaluation methods and the ongoing debate on the potentials and limitations of psychoacoustic methods to assist in sonification research.

PAMPAS, a PsychoAcoustical Method for the Perceptual Analysis of multidimensional Sonification utilizing a the Maximum Likelihood Proceedure in a Just Noticeable Difference experiment experiment has been adapted to be applicable on multi-dimensional sonifications, to achieve linearity during the sonification design stage and to quantify the perceptual resolution, hysteresis effects and perceptual interferences of sonification dimensions and polarities during the evaluation stage. The proposed method helps to design and evaluate single multi-dimensional sonifications, but also to compare multiple sonification designs independently from one another.

The proposed Maximum Likelihood Procedure for the evaluation of two-dimensional sonification designs fulfills the three main quality criteria of scientific test:

1. It is objective, as the procedure and the measures are standardized, and even the qualitative categories are derived from quantitative thresholds.
2. It is reliable, as has been demonstrated by numerous comparative studies Shelton et al. (1982); Pentland (1980); Madigan and Williams (1987) and also has a high test-retest reliability
3. Its validity is high for passive data sonification because experiment participants carry out a task related to typical usage of passive sonification.

Applying the proposed experiment on a two-dimensional sonification of a normalized data space yields quantitative and qualitative results that allow to evaluate a single sonification design. As the method is standardized, the experiment is repeatable and results can be compared between multiple studies.

Note that the proposed method has several limitations.

1. the participants in the experiment are attentively focus on judging the sound. This may also be the case for the intended use case of the sonification. But it may also be that the actual use case involves shared attention, which is likely to increase the JND and the amount of interference, due to cognitive aspects Anderson and Sanderson (2009).

2. the participant in the experiment do not interact with the sound. This may be the case in a data monitoring task. But it may also be that the user interacts with the data and the sound. It can reduce the JND and also hysteresis effects if changes in sound are meaningfully related to (inter-)action (cf. Ziemer and Schultheis (2018)). But it may also increase the JND if the sound does not behave in the expected way.

3. the experiment is unimodal. This may be the same for the intended use case. But in other use cases the user may also receive cues through other sensory systems, such as vision and touch. Multi-modal interference, such as the McGurck effect and the ventriloquist effect (Ziemer, 2020, chap. 4) between vision and audition, may affect the JND, interferences and hystereses.

4. the experiment gives participants sufficient time to listen closely to the sonification. In a real use case the sound may change quickly. The experiment does not reveal the temporal resolution of the multidimensional sonification, i.e., it is possible that the JND is increases when the data fluctuates quickly.

5. participants in the experiment are new to the stimuli. This may certainly be the case for first-time users of the sonification. However, users may familiarize with the sound and thereby learn what to listen for, i.e., which sound aspect to focus on, potentially reducing the JND and interferences. The experiment does not reveal such kinds of training effects, which should be addressed through longitudinal studies Ziemer and Schultheis (2021); **?**.

6. the task in the described experiment is abstract. In an actual implementation of the sonification the sound has a certain context and a relationship to a certain task or aspect related to the sonified data, which may affect the sound perception due to cognitive effects.

7. observations on interferences between two dimensions only focus on small, equal changes of both dimensions, i.e., the diagonal. It is not clear how these interferences transfer to larger changes, and to directions other than the diagonal.

8. The method is limited to aspects of auditory perception and neither includes perception from other senses, nor cognitive aspects of the participants or technical aspects of the user interface. Consequently, it should supplement the well-established methods of sonification design and evaluation, like navigation tasks, NASA Task Load Index and subjective questionnaires.

9. The method is only described for two-dimensional sonifications, as it is typical to test for orthogonality between two dimensions at a time Neuhoff (2004). However, to evaluate a three-dimensional sonification the experiment can be conducted repeatedly, using x-y, x-z and y-z axes Ziemer and Schultheis (2019c).

Despite these shortcomings the proposed method is a good starting point for a perceptually-meaningful sound evaluation during the sonification design stage and in the assessment phase. It should be followed by additional experiments that address the open questions.

Note that many more psychoacoustic methods exist and may be suitable to aid sonification design and evaluation if properly adapted to the aims and signals of sonification. For example consulting psychoacoustic models Zwicker and Fastl (1999), especially those designed for transient sounds like Leman (2000); Sottek (2016), could give a reasonable prediction of perceptual linearity and orthogonality, and *psychoacoustic annoyance* (Zwicker and Fastl, 1999, chap. 16).

7 NOMENCLATURE

I to IV $=$ Quadrants 1 to 4 in a Cartesian coordinate system
JND $=$ **J**ust **N**oticeable **D**ifference
LFO $=$ **L**ow **F**requency **O**scillator
MLP $=$ **M**aximum **L**ikelihood **P**rocedure
NASA-TLX $=$ **NASA** **T**ask **L**oad **I**ndex
PAMPAS $=$ **P**sycho**A**coustic **M**ethod for the **P**erceptual **A**nalysis of multidimensional **S**onification

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

Both authors have conceptualized the described method and tested the procedure. Tim Ziemer has written the initial draft of the manuscript. Holger Schultheis has reviewed and complimented the manuscript.

ACKNOWLEDGMENTS

We thank Rodrigo Gutierrez who has implemented the proposed method in python and Thomas Barkowsky who served as a participant in testing the method.

REFERENCES

Albrecht, R., Väänänen, R., and Lokki, T. (2016). Guided by music: pedestrian and cyclist navigation with route and beacon guidance. *Pers Ubiquit Comput* 20, 121–145. doi:10.1007/s00779-016-0906-z

and, B. B. (2019). Design and evaluation of an audio game-inspired auditory map interface. In *25th International Conference on Auditory Display (ICAD 2019)* (Newcastle upon Tyne, UK), 20–27. doi:10.21785/icad2019.051

Anderson, J. E. and Sanderson, P. (2009). Sonification design for complex work domains: Dimensions and distractors. *Journal of Experimental Psychology: Applied* 15, 183–198. doi:10.1037/a0016329

Asendorf, M., Kienzle, M., Ringe, R., Ahmadi, F., Bhowmik, D., Chen, J., et al. (2021). Tiltification — an accessible app to popularize sonification. In *Proc. 26th International Conference on Auditory Display (ICAD2021)* (Virtual Conference), 184–191. doi:10.21785/icad2021.025

Aures, W. (1985). Berechnungsverfahren für den sensorischen wohlklang beliebiger schallsignale (a model for calculating the sensory euphony of various sounds). *Acustica* 59, 130–141

Barrass, S. (1997). *Auditory Information Design*. Ph.D. thesis, Australian National University, Canberra

Barrass, S. and Kramer, G. (1999). Using sonification. *Multimedia Systems* 7, 23–31. doi:10.1007/s005300050108

Black, D., Hettig, J., Luz, M., Hansen, C., Kikinis, R., and Hahn, H. (2017). Auditory feedback to support image-guided medical needle placement. *International Journal of Computer Assisted Radiology and Surgery* 12, 1655–1663. doi:10.1007/s11548-017-1537-1

Bonebright, T. L. (2001). Perceptual structure of everyday sounds: A multidimensional scaling approach. In *Proc. 7th International Conference on Auditory Display (ICAD2001)* (Espoo, Finland), 73–78

Bonebright, T. L. and Flowers, J. H. (2011). Evaluation of auditory display. In *The Sonification Handbook*, eds. T. Hermann, A. Hunt, and J. G. Neuhoff (Berlin: COST and Logos). 111–144

Bovermann, T., Rohrhuber, J., and de Campo, A. (2011). Laboratory methods for experimental sonification. In *The Sonification Handbook*, eds. T. Hermann, A. Hunter, and J. G. Neuhoff (Berlin: COST and Logos), chap. 10. 237–272

Brewster, S. (2003). Nonspeech auditory output. In *The Human-Computer Interaction Handbook*, eds. J. A. Jacko and A. Sears (London: Lawrence Erlbaum Associates). 220–239

Canévet, G. and Scharf, B. (1990). The loudness of sounds that increase and decrease continuously in level. *The Journal of the Acoustical Society of America* 88, 2136–2142. doi:10.1121/1.400110

Chambers, C. and Pressnitzer, D. (2014). Perceptual hysteresis in the judgment of auditory pitch shift. *Atten Percept Psychophys* 76, 1271–1279. doi:10.3758/s13414-014-0676-5

Degara, N., Kuppanda, T., and Nagel, F. (2013). The walking game: A framework for evaluating sonification methods in blind navigation. In *Proceedings of the 4th Interactive Sonification Workshop (ISon2013)* (Erlangen), 52–57

Fechner, G. T. (1860). *Elemente der Psychophysik* (Leipzig: Breitkopf und Härtel)

Ferguson, S., Cabrera, D., Beilharz, K., and Song, H.-J. (2006). Using psychoacoustical models for information sonification. In *12th International Conference on Auditory Display* (London), 113–120

Fernstrom, M., Brazil, E., and Ottaviani, L. (2003). A new experimental technique for gathering similarity ratings for sounds. In *Proc. 9th International Conference on Auditory Display (ICAD2003)* (Boston, MA, USA), 238–242

Feron, F.-X., Frissen, I., and Guastavino, C. (2009). Upper limits of auditory motion perception: the case of rotating sounds. In *Proceedings of the 15th International Conference on Auditory Display (ICAD2009)* (Copenhagen, Denmark), 6 pages

Fitch, W. and Kramer, G. (1993). Sonifying the body electric: Superiority of an auditory over a visual display in a complex multivariate system. In *Auditory Display: Sonification, Audification and Auditory Interfaces*, ed. G. Kramer (Reading (MA): Perseus Publishing), vol. XVIII of *SFI Studies in the Sciences of Complexity*, 307–325

Grassi, M. and Soranzo, A. (2009). Mlp: a matlab toolbox for rapid and reliable auditory threshold estimations. *Behavior Research Methods* 41, 20–28. doi:10.3758/BRM.41.1.20

Green, D. M. (1990). Stimulus selection in adaptive psychophysical procedures. *The Journal of the Acoustical Society of America* 87, 2662–2674. doi:10.1121/1.399058

Green, D. M. (1993a). A maximum-likelihood method for estimating thresholds in a yes-no task. *The Journal of the Acoustical Society of America* 93, 2096–2105. doi:10.1121/1.406696

Green, D. M. (1993b). A maximum-likelihood method for estimating thresholds in a yes–no task. *The Journal of the Acoustical Society of America* 93, 2096–2105. doi:10.1121/1.406696

Greenwood, D. D. (1997). The mel scale's disqualifying bias and a consistency of pitch-difference equisections in 1956 with equal cochlear distances and equal frequency ratios. *Hearing Research* 103, 199–224. doi:10.1016/S0378-5955(96)00175-X

Greindl, A., Heidegger, P., Groß-Vogt, K., and Weger, M. (2020). Expergefactor: Sonic interaction design for an alarm clock app. In *Proceedings of the 15th International Audio Mostly Conference (AM'20)* (Graz), 4 pages. doi:10.1145/3411109.3411149

Hansen, C., Black, D., Lange, C., Rieber, F., Lamadé, W., Donati, M., et al. (2013). Auditory support for resection guidance in navigated liver surgery. *The International Journal of Medical Robotics and Computer Assisted Surgery* 9, 36–43. doi:10.1002/rcs.1466

Hermann, T. (2002). *Sonification for Exploratory Data Analysis*. Ph.D. thesis, Bielefeld University, Bielefeld

Hermann, T. (2008). Taxonomy and definitions for sonification and auditory display. In *14th International Conference on Auditory Display* (Paris), 8 pages

Hermann, T. (2018). Wave space sonification. In *24th International Conference on Auditory Display (ICAD2018)* (Houghton, MI), 49–56. doi:10.21785/icad2018.026

[Dataset] Hermann, T. (2021). Sonification – a definition. `https://sonification.de/son/definition/`. Accessed: 2021-03-19

Hermann, T., Hildebrandt, T., Langeslag, P., and Rinderle-Ma, S. (2015). Optimizing aesthetics and precision in sonification for peripheral process-monitoring. In *Proceedings of the 21st International Conference on Auditory Display (ICAD2015)* (Graz, Austria), 317–318

Ibrahim, A. A. A., Yassin, F. M., Sura, S., and Andrias, R. M. (2011). Overview of design issues and evaluation of sonification applications. In *Proc. International Conference on User Science and Engineering (i-USEr)* (Selangor, Malaysia), 77–82. doi:10.1109/iUSEr.2011.6150541

Khan, R. A. and Jeon, M. (2018). "musical exercise" for people with visual impairments: A preliminary study with the blindfolded. In *24th International Conference on Auditory Display (ICAD2018)* (Houghton, MI), 204–211. doi:10.21785/icad2018.030

Komatsu, T. and Yamada, S. (2016). Can monaural auditory displays convey directional information to users? In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, eds. A. Papafragou, D. Grodner, D. Mirman, and J. C. Trueswell (Austin, TX: Cognitive Science Society), 930–935

Kramer, G. (1994). Some organizing principles for representing data with sound. In *Auditory Display: Sonification, Audification, and Auditory Interfaces*, ed. G. Kramer (Reading, MA: Addison-Wesley), Santa Fe Studies in the Science of Complexity, Proc. Vol. XVIII. 185–221

Kuppanda, T., Degara, N., Worrall, D., Thoshkahna, B., and Müller, M. (2015). Virtual reality platform for sonification evaluation. In *The 21ts International Conference on Auditory Display (ICAD)* (Graz), 117–124

Leman, M. (2000). Visualization and calculation of the roughness of acoustical musical signals using the synchronization index model (sim). In *Proceedings of the third international conference on digital audio effects (DAFx-00)* (Verona), 6 pages

Lokki, T. and Gröhn, M. (2005). Navigation with auditory cues in a virtual environment. *IEEE MultiMedia* 12, 80–86. doi:10.1109/MMUL.2005.33

Madigan, R. and Williams, D. (1987). Maximum-likelihood psychometric procedures in two-alternative forced-choice: Evaluation and recommendations. *Perception & Psychophysics* 42, 240–249. doi:10.3758/BF03203075

Martens, W. L. (2002). Rapid psychophysical calibration using bisection scaling for individualized control of source elevation in auditory display. In *Proceedings of the 8th International Conference on Auditory Display (ICAD2002)* (Kyoto, Japan), 8 pages

Nagel, F., Stöter, F.-R., Degara, N., Balke, S., and Worrall, D. (2014). Fast and accurate guidance – response times to navigational sounds. In *ICAD* (New York, NY), 5 pages

Neese, M. A. and Walker, B. N. (2009). Auditory interfaces and sonification. In *The Universal Access Handbook*, ed. Stephanidis (New York, NY, USA: CRC Press). 507–522

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature* 395, 123–124

Neuhoff, J. G. (2004). Interacting perceptual dimensions. In *Ecological Psychoacoustics*, ed. J. G. Neuhoff (San Diego (CA): Elsevier), chap. 10. 249–269

Neuhoff, J. G. (2019). Is sonification doomed to fail? In *Proc. 25th International Conference on Auditory Display* (Newcastle upon Tyne), 3 pages. doi:10.21785/icad2019.069

Neuhoff, J. G. and Kramer, G. (2000). Sonification and the interaction of perceptual dimensions: Can the data get lost in the map? In *Proc. 6th International Conference on Auditory Display (ICAD2000)* (Atlanta, GA), 6 pages

Neuhoff, J. G., Kramer, G., and Wayand, J. (2002). Pitch and loudness interact in auditory displays: Can the data get lost in the map? *Journal of Experimental Psychology: Applied* 8, 17–25. doi:10.1037/1076-898X.8.1.17

Neumann, A., Hermann, T., and Tünnermann, R. (2013). Interactive sonification to support joint attention in augmented reality-based cooperation. In *Proceedings of ISon 2013, 4th Interactive Sonification Workshop* (Erlangen, Germany), 58–64

Olivetti Belardinelli, M., Federici, S., Delogu, F., and Palmiero, M. (2009). Sonification of spatial information: Audio-tactile exploration strategies by normal and blind subjects. In *Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments: 5th International Conference, UAHCI 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009. Proceedings, Part II*, ed. C. Stephanidis (Berlin, Heidelberg: Springer). 557–563. doi:10.1007/978-3-642-02710-9_62

Parseihian, G., Gondre, C., Aramaki, M., Ystad, S., and Kronland-Martinet, R. (2016). Comparison and evaluation of sonification strategies for guidance tasks. *IEEE Trans. Multimedia* 18, 674–686. doi:10.1109/TMM.2016.2531978

Pauletto, S. and Hunt, A. (2007). Interacting with sonifications: An evaluation. In *Proc. 13th International Conference on Auditory Display (ICAD2007)* (Montréal, Canada), 519–525

Pentland, A. (1980). Maximum likelihood estimation: The best PEST. *Perception & Psychophysics* 28, 377–379. doi:10.3758/BF03204398

Pires, D., Furtado, B., Carregã, T., Reis, L., Pereira, L. L., Craveirinha, R., et al. (2013). The blindfold soundscape game: A case for participation-centered gameplay experience design and evaluation. In *Proceedings of the 8th Audio Mostly Conference* (New York, NY, USA: Association for Computing Machinery), AM '13, 1–7. doi:10.1145/2544114.2544122

Rönnberg, N. (2019). Sonification supports perception of brightness contrast. *J Multimodal User Interfaces* 13, 373–381. doi:10.1007/s12193-019-00311-0

Scaletti, C. (2018). Sonification $\neq$ music. In *The Oxford Handbook of Algorithmic Music*, eds. R. T. Dean and A. McLean (New York, NY: Oxford University Press). 363–385. doi:10.1093/oxfordhb/9780190226992.013.9

Scheminzky, F. (1943). *Die Welt des Schalls* (Salzburg: Das Bergland-Buch)

Schneider, A. (2018a). Fundamentals. In *Springer Handbook of Systematic Musicology*, ed. R. Bader (Berlin, Heidelberg: Springer). 559–603. doi:10.1007/978-3-662-55004-5_30

Schneider, A. (2018b). Pitch and pitch perception. In *Springer Handbook of Systematic Musicology*, ed. R. Bader (Berlin, Heidelberg: Springer Berlin Heidelberg). 605–685. doi:10.1007/978-3-662-55004-5_31

Seiça, M., Roque, L., Martins, P., and Cardoso, F. A. (2020). Contrasts and similarities between two audio research communities in evaluating auditory artefacts. In *Proceedings of the 15th International Conference on Audio Mostly* (New York, NY, USA), 183–190. doi:10.1145/3411109.3411146

Shelton, B. R., Picardi, M. C., and Green, D. M. (1982). Comparison of three adaptive psychophysical procedures. *The Journal of the Acoustical Society of America* 71, 1527–1533. doi:10.1121/1.387806

Smith, S. (1990). Representing data with sound. In *Proc. IEEE Visualization* (Piscataway (NJ))

Smith, S., Pickett, R. M., and Williams, M. G. (1994). Environments for exploring auditory representations of multidimensional data. In *Auditory display : sonification, audification, and auditory interfaces*, ed. G. Kramer (Reading, MA: Addison-Wesley). 167–183

Sottek, R. (2016). A hearing model approach to time-varying loudness. *Acta Acust. United Ac.* 102, 725–744. doi:10.3813/AAA.918989

Stevens, S. S. and Volkmann, J. (1940). The relation of pitch to frequency: A revised scale. *The American Journal of Psychology* 53, 329–353. doi:10.2307/1417526

Supper, A. (2012). The search for the ”killer application”: Drawing the boundaries around the sanification of scientific data. In *The Oxford Handbook of Sound Studies*, eds. T. Pinch and K. Bijsterveld (New York, NY: Oxford University Press). 249–

Vickers, P. (2006). Lemma 4: Haptic input + auditory display = musical instrument? In *Haptic and Audio Interaction Design*, eds. D. McGookin and S. Brewster (Berlin, Heidelberg: Springer Berlin Heidelberg), 56–67. doi:10.1007/11821731_6

Vogt, K. (2011). A quantitative evaluation approach to sonifications. In *Pro. 17th International Conference on Auditory Display (ICAD2011)* (Budapest, Hungary), 7 pages

Walker, B. N. (2002). Magnitude estimation of conceptual data dimensions for use in sonification. *Journal of Experimental Psychology: Applied* 8, 211–221. doi:10.1037/1076-898X.8.4.211

Walker, B. N. and Kramer, G. (2005). Mappings and metaphors in auditory displays: An experimental assessment. *ACM Trans. Appl. Percept.* 2, 407–412. doi:10.1145/1101530.1101534

Walker, B. N. and Lindsay, J. (2006). Navigation performance with a virtual auditory display: Effects of beacon sound, capture radius, and practice. *Human Factors* 48, 265–278

Walker, B. N. and Nees, M. A. (2011). Theory of sonification. In *The Sonification Handbook*, eds. T. Hermann, A. Hunt, and J. G. Neuhoff (Berlin: COST and Logos), chap. 2. 9–39

Watson, M. and Sanderson, P. (2004). Sonification supports eyes-free respiratory monitoring and task time-sharing. *Human Factors* 46, 497–517. doi:10.1518/hfes.46.3.497.50401

Wegner, C. M. and Karron, D. (1997). Surgical navigation using audio feedback. In *Medicine Meets Virtual Reality: Global Healthcare Grid*, eds. K. Morgan, H. Hoffman, D. Stredney, and S. Weghorst (Washington D.C.: IOS Press, Ohmsha), vol. 39 of *Studies in Health Technology and Informatics*. 450–458. doi:10.3233/978-1-60750-883-0-450

Worrall, D. (2014). Can micro-gestural inflections be used to improve the soniculatory effectiveness of parameter mapping sonifications? *Organised Sound* 19, 52–59. doi:10.1017/S135577181300040X

Worrall, D. (2019). *Sonification Design* (Cham: Springer). doi:10.1007/978-3-030-01497-1

Yang, J. and Hunt, A. (2013). Sonic trainer: Real-time sonification of muscular activity and limb positions in general physical exercise. In *Proceedings of ISon 2013, 4th Interactive Sonification Workshop* (Erlangen, Germany), 44–51

Yeung, E. S. (1980). Pattern recognition by audio representation of multivariate analytical data. *Anal. Chem.* 52, 1120–1123. doi:10.1021/ac50057a028

Ziemer, T. (2020). *Psychoacoustic Music Sound Field Synthesis*, vol. 7 of *Current Research in Systematic Musicology* (Cham: Springer). doi:10.1007/978-3-030-23033-3

Ziemer, T., Nuchprayoon, N., and Schultheis, H. (2020). Psychoacoustic sonification as user interface for human-machine interaction. *International Journal of Informatics Society* 12, 3–16. doi:10.13140/RG.2.2.14342.11848

Ziemer, T. and Schultheis, H. (2018). A psychoacoustic auditory display for navigation. In *24th International Conference on Auditory Displays (ICAD2018)* (Houghton, MI), 136–144. doi:10.21785/icad2018.007

Ziemer, T. and Schultheis, H. (2019a). Psychoacoustic auditory display for navigation: an auditory assistance system for spatial orientation tasks. *J Multimodal User Interfaces* 13, 205–218. doi:10.1007/s12193-018-0282-2

Ziemer, T. and Schultheis, H. (2019b). Psychoacoustical signal processing for three-dimensional sonification. In *25th International Conference on Auditory Displays (ICAD2019)* (Newcastle), 277–284. doi:10.21785/icad2019.018

Ziemer, T. and Schultheis, H. (2019c). Three orthogonal dimensions for psychoacoustic sonification. *preprint*

Ziemer, T. and Schultheis, H. (2020). Linearity, orthogonality, and resolution of psychoacoustic sonification for multidimensional data. *The Journal of the Acoustical Society of America* 148, 2786–2786. doi:10.1121/1.5147752

Ziemer, T. and Schultheis, H. (2021). The CURAT sonification game: Gamification for remote sonification evaluation. In *26th International Conference on Auditory Display (ICAD2021)* (Virtual conference), 233–240. doi:10.21785/icad2021.026

Ziemer, T., Schultheis, H., Black, D., and Kikinis, R. (2018). Psychoacoustical interactive sonification for short range navigation. *Acta Acustica united with Acustica* 104, 1075–1093. doi:10.3813/AAA.919273

Zwicker, E. and Fastl, H. (1999). *Psychoacoustics. Facts and Models* (Berlin, Heidelberg: Springer), second updated edn. doi:10.1007/978-3-662-09562-1