Using Conceptual Metaphors to Represent Temporal Context in Time Series Data Sonification

Article *in* Interacting with Computers · February 2020

DOI: 10.1093/iwc/iwz036

1 author:

Stephen Roddy
University College Cork

Using Conceptual Metaphors to Represent Temporal Context in Time Series Data Sonification.

ABSTRACT

This article explores how conceptual metaphor theory can be applied to the problem of representing temporal context in the sonification of time series data. It opens with an introduction to some of the conceptual metaphors involved in our understanding of time and music. Two of these metaphors are extended to the domain of Auditory Display and act as a guide to the creation of a design framework for representing temporal context in the sonification of time series data. An empirical listener evaluation of this framework is then presented and the results analysed to gauge the efficacy of this approach. The results suggest that the framework supports the listener, to a limited degree, in estimating which time-point they are currently hearing, how many time-points they have already heard, and how many time-points they have left to hear. Improvements to the framework are suggested and plans for future iterations are discussed.

KEYWORDS

Auditory Display; Sonification; Time Series Data; Embodied Cognition; Conceptual Metaphor

RESEARCH HIGHLIGHTS

- Conceptual Metaphors in Auditory Display Design
- Framework for representing Temporal Context to User/Listener
- Empirical Evaluation indicates the framework provides a limited sense of temporal context.
- It helps listeners to estimate:
  1. Which time-point they are currently hearing,
  2. How many time-points they have already heard
  3. How many time-points they have left to hear

CCS CONCEPTS

Human-centered computing → Auditory feedback
Computing methodologies → Cognitive science
Human-centered computing → Sound-based input / output
Applied computing → Sound and music computing

1. Introduction

Auditory display is a human-computer interaction (HCI) discipline that focuses on displaying information to a user through sound. Sonification is an auditory display technique in which data is mapped to sound for the purposes of representing the original data source and/or making meaning from the data. It is often used with complex datasets which are difficult to interpret through visual means alone, for example, epileptic seizure detection in EEG data (Razavi et al. 2015). Sonification is becoming an increasingly popular tool for helping to represent network activity (Worrall, 2015; Wolf and Fiebrink, 2013; Fairfax et al, 2014) and according to Rimland et al. (2013), this trend is set to continue with the global proliferation of Internet of Things (IoT) technologies. However, making sense of complex IoT data which can span over large ranges of time requires some more basic research in techniques for representing time to an end user. As Flowers (2005) points out, representing time series data with sound can be problematic. When attempting to do so, he argues that "judicious selection of variables and mappings can be an issue; plotting “everything at once” in a dense multivariate data sample can lead to “a mess” in either sound or sight". Instead, Flowers argues for sequential comparisons of time series data while allowing time to "represent itself". This involves the organization of sounds over time to reflect relationships in the original time series rather than representing time points with specific sounds. This approach is explored in greater depth in this article.

Third Wave HCI has been defined in part by the adoption of design principles informed by research in the discipline of embodied cognition (Dourish 2004) as researchers "embraced experience and meaning-making" (Bødker 2015). A number of researchers have adopted principles inspired by theories developed in embodied cognitive science research: image schemata, conceptual metaphors and conceptual blends, in the design of user-friendly solutions for HCI problems (see Imaz and Benyon, 2007; Hurtienne, 2009; Macaranas et al., 2012; Waterworth and Riva, 2014; Bødker and Klokmose 2016). A similar approach has also been adopted in auditory display where researchers have drawn from these theories to design more user-friendly systems (see Antle et al., 2011; Brazil & Fernström, 2006; Diniz et al., 2012). Correspondingly, our design framework draws on Lakoff and Johnson's (1980) conceptual metaphor theory to help give the sonification listener a sense of temporal context.

2.2 Temporal Context and Conceptual Metaphors

The expression "temporal context" is used in this article to reference both the listener's understanding of where they are in a sonification in regards to its start and end points and where they are in a time series in regards to what time-points they are currently hearing, have previously heard and will hear before the sonification ends. The representation of temporal context to a listener can be problematic in sonification. Visual graphing is not time bound. At any given time the entire time series in a graph is accessible to an observer. Visual graphs present static information to an observer who can choose where in the graph to direct their attention. Auditory displays evolve and unfold as time-bound experiences rather than static objects. They must dynamically present their information and as a result, the listener does not have the luxury of directing her/his attention to sets of features that unchangingly endure through time. This presents some issues for time series sonification where a sense of temporal context can be key to interpreting the data. When attending to a well designed visual graph of time series data, the observer can clearly identify individual time-points and temporal relationships, on the basis of context cues like axis markings, and labels (Hermann, 2002, p. 12). It has been argued that auditory context cues are also required for a listener to interpret an auditory graph (Walker and Nees, 2011). The design framework described in this article builds on these points to create a sense of temporal context through the spatial organisation of the sonification. It is intended for use in situations where the designer knows the length of the time series of sonified data.

[TABLE 1]
Table 1. The Moving Times Metaphor

Lakoff and Johnson (1980) introduced the concept of a conceptual metaphor as a function of cognition that allows us to make sense of one conceptual domain in terms of another. They argue that conceptual metaphor is an active process at the level of the "cognitive unconscious" which supports reasoning and imagination. It makes abstract and unfamiliar domains of perceptual and conceptual experience meaningful by framing them in terms of more familiar domains (Lakoff & Johnson, 1980, 1999; Lakoff & Núñez, 2000). This is achieved by the mapping of roles from familiar areas of experience, termed source domains, onto corresponding roles in unfamiliar target domains. In this way, unfamiliar or abstract target domains that might otherwise remain meaningless or unknowable can be interpreted and understood in terms of the more familiar source domain (for further details see Lakoff and Johnson, 1980). The classic example of a conceptual metaphor is the *love is a journey* metaphor in which one's knowledge about and previous experiences of journeying are mapped to lend a familiar structure to the abstract concept of love. This allows love to be conceptualised using the logical structure inherent to the concept of a journey. This involves a beginning, middle, and end where the lovers are travellers on a common path along which they may encounter difficulties and perils just as one might expect in a journey. Imaz and Benyon (2007) demonstrate how conceptual metaphors can be applied to HCI and software engineering to create intuitive user centred designs and Roddy and Bridges (2018a) expand upon this to describe how conceptual metaphors can be used in the design of intuitive auditory display solutions.

[TABLE 2]
Table 2. The Moving Observer Metaphor

The design framework discussed here borrows from the work of Cox (1999) and Johnson and Larson (2003), which will be discussed shortly. They, in turn, build upon Johnson's earlier work with Lakoff (Lakoff and Johnson, 1980) on the metaphorical conceptualisation of time. These researchers argue that time is conceptualised metaphorically as movement through space. This means that people frame, understand and reason about the abstract concept of time in terms of the familiar concepts of space and spatial movement. For example, Johnson and Larson (2003) point out that words used to speak about the passage of time, e.g., passage, flying, crawling, coming, going, front, behind, are also used to talk about movement through space.
Two chief systems of metaphor are thought to define how time is understood and reasoned about (Cox, 1999; Johnson and Larson, 2003; Lakoff and Johnson, 1980). These are the *moving times* metaphor presented in table 1 and the *moving observer* metaphor presented in table 2. According to Johnson and Larson (2003):

> In the *moving times* metaphor, the times are the figure moving relative to the stationary observer (as ground), whereas in the *moving observer* metaphor, the observer is the figure moving relative to the time landscape (as ground).
>
> Johnson and Larson (2003).

In both systems, future times approach the observer's location from the front, become present when they share a location with the observer, and fade into the past when they pass behind the observer's location. The basic cross-domain mappings involved in these metaphorical systems, as documented by Johnson and Larson (2003), are presented in tables 1 and 2. Both systems are thought to play a key role in the observer's understanding of time.

[TABLE 3]
Table 3. Conceptual Metaphor Guide

Johnson and Larson (2003) extend this metaphorical conception of time as spatial movement to music, arguing that listeners understand and reason about music as if it were spatial movement also. This argument is based around the three core systems of metaphorical mappings of *moving music*, *musical landscape* and *music as moving force*. These metaphors draw from the listener's perceptual experience of objects in motion within a spatial field, self-determined motion through spatial fields, and forced motion through spatial fields. In the *moving music* metaphor, musical events are entities that move from a point in front of the listener (future) past the listener's position (present) to a point behind the listener (past). In this metaphor, musical events are in motion while the listener remains static. In the *musical landscape* metaphor music is a landscape of events that a listener must navigate through. The area in front of the listener is the future, the area behind is the past and their current position is the present. Here the listener is moving while the musical events remain static. The *musical landscape* metaphor suggests a complex musical soundscape that changes and evolves as the observer moves through it experiencing the different sounds generated in different areas. In the *music as a moving force* metaphor, a listener who might otherwise remain static is being pushed or forced through a series of locations populated by musical events. In this metaphor, the listener has no agency or power over her/his own trajectory of motion. Sonification and music involve qualitatively different listening modes (Vickers 2012) and as a result, we would not necessarily expect the conceptual metaphors underpinning musical listening to be directly relevant to sonification. Nevertheless, the "*moving observer*" and "*moving times*" metaphors, are thought to be crucial to our understanding of time more generally. As such we would expect the conceptual metaphors involved in the interpretation of sonified time series data to share similar features to these more basic metaphors for time.

2.3 Towards Design Principles

As mentioned previously, conceptual metaphors have been applied to aid in the design of user-centric HCI systems (Imaz and Benyon, 2007; Hurtienne, 2009; Macaranas et al., 2012) and auditory displays (Walker and Kramer, 1996). Antle et al (2009) have demonstrated how conceptual metaphors can inform sonification mapping strategies for user-centric auditory displays across three distinct research projects. This process generally involves determining which conceptual metaphors might be involved in the listener's interpretation of the display and designing data to sound mapping strategies that reflect those metaphors. This in turns helps the user to correctly interpret the system. Similarly, Roddy and Bridges (2018b) present a sonification mapping strategy for IoT networks informed by a popular metaphor for machine to machine (m2m) communication, that of "machines talking to each other". Given that the conceptual metaphors discussed here are critically important to our understanding of time it may be possible to draw from them to create a sonification design framework for the representation of time series data.
In the original *moving times* and *moving observer* metaphors the future maps onto the area in front of the listener, the present maps onto the listener's location and the past maps onto the area behind the listener. These mappings also hold true for all three of the musical metaphors. Furthermore, the *moving times* and *moving observer* metaphors are posited as our basic metaphors for understanding and interpreting concepts related to time. Given both of these points, we might expect that this mapping of temporal concepts onto spatial parameters might also hold for our understanding of time series data in a sonification context. These mappings, and how they might be represented sonically, are represented in table 3 and were derived from the original metaphors as follows:

In the *moving times* metaphor, individual times are mapped onto objects and the perceived motion of those objects represents the passage of time. Time series sonification generally involves the presentation of data to a listener using a sequence of time-indexed sonic events where each individual event represents an individual data point. Accordingly, we might expect listeners to interpret time-indexed data points as if they were physical objects and to interpret changes in the data as changes undergone by those physical objects. The metaphors discussed previously provide no mapping for motion in the vertical plane. We might expect that changes in data value could be mapped to the vertical path of motion for an object moving through space. This is motivated by the previous mappings of data points to physical objects and changes in the data to changes in physical objects. Likewise, while the metaphors prescribe no mappings for the horizontal plane, we might expect that this plane can be treated in a similar manner to visual graphs in which time series data is organised at set intervals, or units of time, along a left to right axis. The *moving times* and *moving observer* metaphors map the passage of time to front/back movement, and this might also be exploited to represent these changes in the data over set intervals or units of time. Another factor in these metaphors is distance. In both metaphors increased distance in terms of physical space maps to increased distance in time and vice versa. This can be exploited to help organise the framework.

From a sonic point of view, listeners tend to relate pitch with vertical height (Rusconi et al., 2006) and this has been widely exploited by auditory display designers (Cabrera and Tilley, 2003). As such pitch might present a useful parameter for representing the vertical plane in the framework. Distribution of sonified data across the horizontal plane of the stereo field has also proved to be an effective organisational principle in auditory display contexts (Walker and Nees, 2011) and can be employed again here. Perceived distance from the listener's front and back might be controlled to structure the presentation of data over a single unit of time, while distance form the listener's left and right might be controlled to structure the presentation of the entire time series. However, there are a number of perceptual and psychoacoustic constraints that need to be accounted for when developing a design framework on the basis of the principles discussed thus far.

2.4 Perceptual and Psychoacoustic Constraints

The topic of binaural sound localisation, the mechanisms, and processes by which the ear and brain resolve the spatial location of a sound is complex. Generally, we rely on interaural time difference (ITD), interaural level difference (ILD) and spectral filtering by the pinnae and head (HRTF) to resolve locations (Middlebrooks & Green, 1991). However, these processes interact in a sophisticated manner and their operation is further complicated when the effects of cues from other sensory modalities, such as vision, are taken into account. The framework proposed in this article is designed for use with a 2-channel stereo audio format. With the standard stereo setup, the listener generally perceives the full stereo field as occupying a roughly 60-degree frontal angle, although this can be expanded to 180-degrees when cross-feed is cancelled over headphone presentation. While this is a less than ideal setup stereo has nonetheless become something of a ubiquitous format. The stereo format doesn't directly allow for the direct representation of verticality (azimuth) although this can be achieved through binaural and ambisonic recording techniques, with the former requiring headphone presentation or crosstalk cancellation. The framework described here bypasses this problem somewhat by using pitch to represent vertical height. However, this introduces the issue of frequencey/pitch discrimination. The just-noticeable difference (JND) for frequency is highly dependant on the frequency content of the sound heard, and varies across the frequency spectrum from roughly 3hz for frequencies below 500hz to .6% for those above 1000hz (Benesty et al., 2007). The average human can perceive roughly 1400 discrete frequencies (Olson, 1967) but can only make absolute judgements of (or correctly identify) up to 6 tones, though musically trained listeners with perfect pitch might discern between 50 and 60 (Miller, 1956). This suggests that for the musically untrained listener the pitch aspect of the framework will be somewhat more difficult to interpret.

Front/back localisation presents another challenge. In general, front/back localisation can be somewhat ambiguous because when a source is encountered head on there should be no perceptible ITD or ILD to aid in the localisation process. This results in a "cone of confusion" in which the front/back orientation becomes all but impossible to determine over the cone-shaped set of points projected laterally from a listener's interaural midpoint (Woodworth, 1938). In a natural environment this problem can be addressed by shifting the direction of the head and pinnae, gaining the spectral cues that are critical to both vertical and front/back source localisation. But, because front/back cues are reliant on this spectral filtering by the head and pinnae, it can be difficult to represent front/back orientation for a listener in a stereo-field (Middlebrooks & Green, 1991). This presents a problem for the framework presented here as front/back movement is a critical aspect of the conceptual metaphors that have informed its design.
Another issue to be considered is minimum audible angle (MAA). MAA is the smallest angular separation of sound sources that is discernible by a listener. It is highly dependant on both the frequency of the sound and the angle of presentation and as such varies widely and in a non-linear manner (Mills, 1958; Perrott and Saberi, 1990). MAA needs to be accounted for when distributing sounds across the stereo field, to ensure that listeners can differentiate between angles of presentation.

The ability of the auditory system to discriminate the distance of a sound source is substantially less accurate than its discrimination of vertical and horizontal angle (Middlebrooks & Green 1991). The main cues for distance perception are intensity, ratio of direct-to-reverberant (D/R) energy ratio, and ILD, but distance perception is also influenced by familiarity of the sound source, source sound pressure level, source spectrum, radiation pattern, relative azimuth, environmental acoustics, and near-versus-far range of distances (Kolarik et al., 2016). This is further complicated the influence of visual cues on distance perception and this complex interaction of factors can make distance perception somewhat difficult to control by artificial means.

An additional factor to be taken into account is echoic memory. This is the form of sensory memory that stores the sounds that a listener has just heard. It operates by holding sounds in an unprocessed state until a subsequent sound is presented in relation to which the original sound becomes meaningful (Clarke, 1987). This holding operation is estimated to have a capacity of roughly 4 seconds (Darwin et al., 1972). Working memory is responsible for the short-term holding and processing of information and is often said to have a capacity of 7 ± 2 chunks, or discrete objects, of information (Miller, 1956) and is thought to decay after 10-20 seconds (Cowan, 1988; 1995; 2001). This in turns places some hard limits on the proposed design framework in terms of durations of its component sounds. While care was taken to address the majority of the limits discussed here some of these factors were not fully accounted for in the initial prototype framework. These points are discussed in greater detail in section 5.2.

3. Design Framework.

The proposed design framework operates as follows: A single stream of time series data contains two components, the observed data values, and their associated time points. The number of possible measurement systems by which observed data might be measured or encoded is necessarily quite large and diverse. In contrast, temporal data is routinely coded using standardised units of time (e.g. seconds, days, years, etc.). This provides something of a baseline standard with which we can organise our sonifications. With this framework, the designer chooses a sonic symbol with which to represent their data. This sonic symbol could be generated using sound synthesis techniques or it might involve a more complex process such as that used in the creation of parameterised auditory icons (Gaver 1989). Observed data values are then mapped to modulate salient auditory parameters of that sonic symbol (e.g., pitch, timbre, tempo) resulting in a “data object”. Temporal data are represented in the organisation of these data objects in stereo space in accordance with rules derived from the systems of conceptual metaphor discussed previously. These data objects can be formalised in terms of Rohrhuber’s (2010) sonification variables.

[FIGURE 1]
Figure 1. Unit Sequence from Front to Back

Time is represented in two ways with this framework. The entire time series of the given data set is distributed from left to right along the horizontal axis of the stereo field. This involves the panning of sounds into different bins across the stereo field. Each bin represents a specific “unit sequence” which is a short temporal sequence defined by the designer. The unit sequence is described in greater detail shortly. The Doppler effect is then used to denote the passage of time within each individual unit sequence as illustrated in figure 1. The Doppler effect, illustrated in figure 3, describes the phenomenon where a sound is perceived as higher in frequency when approaching a listener but lower in frequency when retreating from a listener. As a sound source approaches a listener, the crest of each successive sound is emitted at a location that is closer to the listener and this bunching up of sound waves is perceived as an increase in the frequency of the sound. Conversely, as a sound source retreats from a listener successive sound waves are emitted from points of increasing distance corresponding to a drop in the perceived frequency for the listener. The task of representing front and back orientation in stereo-space is a complex one that is not yet solved. However, the Doppler effect leverages the listener’s knowledge of physical sound sources in motion to achieve this in a limited way for moving sound sources. For each individual unit sequence, the Doppler shift will take place at the mid-point of the temporal sequence.

[FIGURE 2]
Figure 2. Time Series from Left to Right

The listener retains a static listening position throughout a sonification while the data objects representing the data move. The first data object in a given data set will sound on the extreme left of the listener’s position at an angle of -90°. Over the course of a sonification, each consecutive data object will sound in sequence at increments of x° from the position of the initial data object. The midpoint of the dataset will sound at the listener’s position 0°. The size of x°, and therefore the number of unique spatial positions that a listener can discern across the stereo field, is determined by how the designer defines the length of a single unit sequence. This is further constrained by the psychoacoustic factors discussed in this article. The last data object will sound to the extreme right of the listener at an angle of +90°. Data objects are distributed across stereo-space by panning them between stereo sources and applying reverb modelling to simulate distance. Gerzon (1992) illustrated how the ratio of reverberated to direct sound, pre-delay time and early reflection amplitude can all be manipulated to create a perceptible sense of distance from the listener’s position. However, it must be noted that the auditory system does not operate with sufficient resolution to allow us to precisely represent the specific distal locations of a sound source to a listener using this method. Rather, these reverb parameters impart a general sense that a sound source is changing in distance. These reverb parameters are manipulated, to make the data objects sound as though they approach the listener from the left and then retreat from the listener on the right. A ratio between direct and reverberated sound that is approaching 0:1, a pre-delay of approaching 100ms and an early reflection amplitude approaching -30db can be used to create a sense of increased distance. A direct to reverberated sound ratio which is approaching 1:0 with a pre-delay time approaching 0ms and an amplitude level approaching 0db gives the impression of decreased distance from the sound source. While this is a somewhat coarse solution to the problem it has the advantage of being implementable in the stereo field. This configuration, represented in figure 2, creates a matrix of vectors that pass by the listener’s position in sequential order moving from left to right across the stereo field.

[FIGURE 3]
Figure 3. Doppler Effect

A unit sequence should contain a meaningful temporal sequence chosen by the designer to strike a balance between the clear and effective presentation of observed data values on one hand and their associated time points on the other. For example, a data set that represents temperature changes in a city sampled daily over the course of a year can be represented with a unit sequence length of 1 month so that 365 samples can be represented in just 12 positions along the horizontal axis of the stereo field. Rather than simply averaging and presenting the data for each month, the concept of a unit sequence allows uniform portions of the data to be presented sequentially while retaining a sense of temporal context for the listener. A single unit sequence does not necessarily need to be presented in its entirety before the subsequent unit sequence is presented. A designer might allow unit sequences to overlap with one another to some degree to aid listeners in comparing sequences to one another. Overlap might support listeners in comparing the unit sequences to one another by shortening the time between presentation and allowing listeners to hear the beginning of one sequence and the ending of another simultaneously.

Using this framework a listener should be able to effectively estimate the value of the current time point heard and the temporal position of the time point in relation to beginning and end points. As the framework is focused on the representation of the time series, the question of which data-values relate to which time-points is of critical importance. As referenced previously Flowers (2005) points out that sonifications containing short individual data streams that can be easily cross-compared and are presented in a sequential manner are more intelligible than continuously represented data-streams. The proposed framework is intended to help listeners estimate the relationship between data-points and time-points and also the relationships between successive time-points presented in sequence.

4.1 Listener Evaluation

4.1.1 Introduction & Participants

A listener evaluation was used to test the effectiveness of the design framework. Data were sonified to the specification of the framework and participants were then asked to listen to the data and answer some questions. Parameter mapping sonification (PMSon) maps data directly to sonic parameters such as pitch and timbre (Grond and Berger, 2011). This was chosen to map data to pitch creating sonic symbols. In this sonification monthly rainfall measured at Dublin Airport over 2014 is mapped to the frequencey of a sine tone. The dataset[1] provided by MET Éireann is presented in table 4. Data objects were generated in Csound and Logic Pro X. The dataset was mapped to control the frequencey of the output of a poscil opcode in Csound. The poscil opcode, a contraction of "precise oscillator", offers fine-grained control of frequency values. It was used to synthesise a simple sine wave signal. Rainfall data was mapped to modulate the frequency of this sine wave across a range from 220Hz (A3) to 1760Hz (A6). The code for this implementation is presented in the appendix. This resulted in a continuous sonification that was then imported into Logic Pro X where the continuous sonification was divided into unit sequences. The Csound code is provided in the supplementary materials for this article. The output of the Csound code and the finished sonification produced in Logic Pro X are also provided.

[TABLE 4]

Table 4. Rainfall in millimetres recorded at Dublin Airport over 2014

Data objects were distributed horizontally across the stereo field by panning them between stereo sources in Logic Pro X and applying reverb modelling to simulate distance. The parameters used in this reverb modelling were informed by the work of Gerzon (1992) discussed previously and are presented in table 5. Each individual data object is 9 seconds long and represents the change in value from one month to the next. There was an overlap of 1/3 of a unit sequence between data objects so that the last 3 seconds of one unit sequence would overlap with the first second of the subsequent unit sequence. As such a new unit sequence would begin to play every 6 seconds. The entire sonification had a duration of 1 minute and 16 seconds. The first data object represents the change in value from 101.6mm of rain for January to 88.5mm for February. A Doppler shift algorithm is applied to each data object to simulate a Doppler shift at the halfway point, 4.5 seconds, in the sound's playback. In order to expand the spectral content of the data objects, and so make the Doppler effect more pronounced, reverb was added at a ratio of 10:1 between direct and reverberated sound. The Waves Doppler plugin was used to simulate Doppler effect. The frequency of the input note is increased by a factor of approximately +1.25. At the shift point, the frequency is reduced by a factor of approximately -1.25 over the course of 350ms. In order to exaggerate this effect amplitude modelling is applied. As the pitch shift approaches the overall amplitude increases by +35db, in a logarithmic fashion. Amplitude becomes loudest at the point of the pitch shift and then begins to fade by -35db over the remainder.

We have chosen to map data values to pitch and then to modulate pitch with a Doppler effect. This design decision was informed by Fitch and Kramer's (1994) 'piggy-back' concept in which aspects of a single sound are modulated by a number of different variables. While the introduction of pitch shifting is likely to complicate the listeners' interpretation of changes in the dataset, it was thought that the sequential presentation of data objects might help to correct for this. As a data object approaches the listener its pitch will be higher. As it retreats after the Doppler shift its pitch will be lower. Regardless, the overall direction of the pitch contour will remain the same. The issue for the listener is to determine if the data presented in the second half of the unit sequence is higher or lower than the data presented in the first half. This question only arises when the direction of the pitch contour reverses. However, the next

[1] The data is available at http://www.met.ie/climate/monthly-data.asp?Num=532

data object in the sequence departs from the value reached at the end of the previous data object. This initial value is presented at the higher pitch level, as the Doppler effect has not yet been applied. The listener can compare this initial pitch value back to the values presented at the start of the previous data object to get a sense of the degree to which the data values have changed.

[TABLE 5]
Table 5. Control Parameters for Left to Right Distribution

The evaluation aimed to test how effectively the listeners could estimate the following when listening to stimuli designed with the design framework:

1. Value of current time-point heard.
2. Temporal position in relation to beginning and end points.
3. Relationship between data-point and time-point.
4. Relationships between data-points presented in temporal sequence.

An online crowdsourcing approach was adopted for the evaluation. Oh and Wang (2012) analyse recent research into the limitations of crowdsourcing for empirical testing to determine the factors which influence the quality of crowdsourced data for music perception experiments. Participants were recruited from 49 countries via online crowdsourcing platform Crowd Flower. These were selected from a pool of Crowd Flower's most trusted participants. Experimental materials were hosted on Survey Gizmo. Participants were compensated financially for their participation. 191 people took part in the evaluations. Of that number 72% were male and 28% were female. 18% of participants were between 18 and 24. 43% of participants were between the ages of 25 and 34. 35% of participants were between the ages of 35 and 54 and 3% of participants were over 55. Precautionary measures were taken to ensure that participants were using proper equipment. Listeners who could not pass a simple evaluation proving that they were using a 2-channel stereo playback were prevented from participating in the evaluation. Listeners completed a short training trial to familiarise themselves with the procedure at the outset of the evaluation. The experiment took approximately 25 minutes to complete. Restrictions on timing were used to ensure that listeners did not spend longer than 30 minutes and less than 20 on the evaluation.

4.1.2 Design, Materials & Procedure

4.1.2.1 Training Phase

At the beginning of evaluation A, the concept of a sonification is explained to the listener. They are given the following information:

"These evaluations deal with a type of audio called sonification. A sonification represents data using sound. It is like a sonic version of a visual graph e.g. A bar chart, pie chart or trend graph. This sonification presents the level of rain in Dublin for each month of 2014.

The rainfall measurement for each month is represented by a sound that flies by the listeners head. When the level of rainfall is high, the sound has a high pitch and when the level of rainfall is low it has a low pitch.

The rainfall measurements for the first half of the year, from January to June, are presented on the listener's left-hand side. January's rainfall passes to the far left of the listener and each month after passes a little nearer to the listener. The rainfall for June passes very close to the listener on the left-hand side.

The rainfall for the second half of the year, from July to December, passes the listeners right-hand side. The rainfall for July passes very close to the listeners right-hand side and each month after passes by a little further from the listener. The rainfall for December passes to the far right of the listener."

Listeners are then presented with this information again and asked: "Please make sure you have read and understood the description above before continuing with the training session.". They must check a box that reads "I have read and understand this description" before moving forward. If they do not they are not allowed to move forward with testing. Those who tick the box then undertake a training evaluation to familiarise them with the kinds of questions that will be asked of them. This evaluation page presents the listener with the previous description of what a sonification is and how this particular sonification is to be interpreted a third time. When a participant finished the training session they began with evaluation A. The audio stimuli for the training session are organised on the basis of the framework but use different data values to the stimuli presented in the main evaluations.

4.1.2.2 Evaluation A

After the training session listeners were presented with five individual evaluations A to E. Each evaluation had an associated audio stimulus. Listeners were allowed to replay the sonification stimuli as many times as was needed to answer the questions. Evaluation A contained the audio for the entire rainfall sonification described in section 4.11. It had a duration of 1 minute and 16 seconds. Each unit sequence was 9 seconds long and subsequent unit sequences were sounded every 6 seconds. The last 3 seconds of one unit sequence would overlap with subsequent unit sequence. The evaluation contained 2 questions. In questions 1 and 2 listeners were asked to listen to the audio stimulus and to answer "In which month was the rainfall at its HIGHEST level?" and "In which month was the rainfall at its LOWEST level?" respectively. Listeners could choose from a list including all 12 months and an additional option of "Don't Know". Evaluation A presents participants with the following instructions:

"The following audio clip is the sonification described on the previous page. Please listen to it fully and answer the questions that follow. If you do not know the answer to a question please select 'Don't Know'.

A reminder of the information about the sonification is provided below:

[The description in section 4.1.2.1 is included here again]

Please listen to the clip and answer the questions that follow. You can listen to the clip as often as needed to help in answering the questions. If you do not know the answer to a question please select 'Don't Know'."

4.1.2.3 Evaluation B

Evaluation B contained questions 3, 4, 5 and 6. The audio for evaluation B contained the sonified data for January, February, and March only and had a duration of 21 seconds. Each individual unit sequence lasted 7 seconds with the last 2.33 seconds of one unit sequence overlapping with the first 2.33 seconds of the following unit sequence. In questions 3 listeners were asked "Which half of the year did the sonification stop playing in?" and given an option of "First Half", "Second Half" or "Don't Know". In questions 4 listeners were asked "What was the last month of the data played?" and given the option of choosing from a list of 12 months with a 13th option of "Don't Know". For question 5 listeners were asked to listen to the audio and answer "How many months of data were played?". The options provided included a list of numbers from 1 to 12 and "Don't Know". For question 6 they were given the same set of possible responses and asked to answer, "How many months of the data were left to play?"

The information and instructions for evaluation B were as follows:

- *The following audio clip is taken from the sonification presented in Evaluation A.*
- *It represents the rainfall values for SOME of the MONTHS in 2014.*
- *Please listen to the clip and answer the questions that follow.*
- *If you do not know the answer to a question please select 'Don't Know'*
- *You may replay to the clip as many times as is required to help answer the questions.*

4.1.2.4 Evaluation C

For evaluation C, listeners were provided with a new audio stimulus, which contained the sonified data from January to September and had a duration of 56 seconds. As with evaluation B, each individual unit sequence lasted 7 seconds with the last 2.33 seconds of one unit sequence overlapping with the first 2.33 seconds of the following unit sequence. Listeners were presented with the same questions they had been asked in the previous evaluation (B). They were asked to listen to the audio and answer the questions. As such, questions 7, 8, 9, and 10 in evaluation C corresponded respectively with questions 3, 4, 5 and 6 from evaluation B. The instructions for evaluation C were the same as B.

4.1.2.5 Evaluations D & E

Evaluation D and E contain a single question each, numbers 11 and 12. Audio D contained the sonified data for August and September and had a duration of 14 seconds while audio E contained the sonified data for April and May and also had a duration of 14 seconds. Each unit sequence lasted 7 seconds with the last 2.33 seconds of one unit sequence overlapping with the first 2.33 seconds of the following unit sequence. Questions 11 and 12 both asked "In which month is the level of rainfall highest?" and the possible answers for both questions were "First Month", "Second

Month", and "Don't Know". The instructions for D and E were almost the same as for B but that the second point of information read: *"It represents the rainfall values for TWO MONTHS in 2014."*

4.1.3 Rationale for the Evaluation Questions

Questions 1 and 2 in evaluation A were intended to determine how effectively listeners could estimate the data values associated with specific time-points. Questions 3 and 4 in evaluation B and questions 7 and 8 in evaluation C were intended to determine how well listeners could determine the specific time point being heard. Questions 5 and 6 in evaluation B and questions 9 and 10 in evaluation C were intended to determine how well listeners could estimate their temporal position in relation to beginning and end points of the sonification. Questions 11 and 12 in evaluations D and E were intended to test how well listeners determine the relationships between sequentially presented data values using this framework. Each question and the rationale behind it is presented in table 6 below.

[TABLE 6]
Table 6. Questions and Rationale

5.1 Results & Analysis

The listener responses for each question were analysed using two chi-square goodness-of-fit tests. The results are presented in table 7. This analysis method was chosen to help determine whether or not the frequency distribution of listeners' responses deviated from the mean expected responses to a statistically significant degree. The first test is performed on the results for all of the options in the evaluations including the 'don't know' options. This gives a sense of the overall statistical significance and practical significance of the data, as indicated by the effect size measure: Cramer's V. The second test is performed on all of the results excluding the 'don't know' options. Performing both of these tests allows for the identification of cases where statistical significance and effect size are caused by the number listeners who chose this option, indicating that they did not understand the stimuli and/or question asked of them.

[TABLE 7]
Table 7. Analyses of Results

For question 1 and question 2 listeners were presented with the entire sonification and asked to determine which months had the highest and lowest rainfall respectively. These results are illustrated in figures 4 and 5 respectively.

[FIGURE 4]
Figure 4. Data Breakdown for Question 1

The high point of the data was in August, but only 15 listeners identified this from the sonification. February was selected with the highest frequency of 23 listeners. This was followed by July with 20 listeners and then October with 17. This suggests that listeners could not effectively link data-values with time-points. 24 listeners selected the 'don't know' option. This was roughly the same number that selected the correct option.

Two chi-square goodness-of-fit tests were carried out on the results. The first test included the results for each of the month options and the 'don't know' option $\chi2$ (12, 191)=21.42, p<.05, φc=.1. The second test considered only the results for the month options $\chi2$ (11, 167)=15.87, *p*>.05, φc=.09. The first test produced a highly significant p-value of p<.001 but a small effect size of φc=.1. This suggests that the results are of high statistical significance but little practical significance. The second test generated a non-significant p-value of *p*>.05 and a small effect size of φc=.09. This further suggests that the number of listeners choosing the correct option over the incorrect alternatives was statistically non-significant and of almost no practical significance.

The results for question 1 suggest a substantial number of listeners misinterpreted the sonification by choosing incorrect options and a similarly substantial number of listeners reported that they did not understand the sonification. This indicates that listeners could not effectively link the highest data point to its time point using the sonification.

[FIGURE 5]
Figure 5. Data Breakdown for Question 2

In question two, the low point of the data was in April, but only 15 listeners identified this from the sonification. June was selected by the highest number of listeners 29, followed by July with 21 listeners and March with 18. Furthermore, 27 listeners reported that they did not know the correct answer. This suggests that listeners could not effectively link data-values with time-points.
Two chi-square goodness-of-fit tests were carried out on the results. The first test included the results for each of the month options and the 'don't know' option $\chi2$ (12, 191)=43.74, *p*<.001, φc=.14. The second test considered only the

results for the twelve-month options $\chi 2$ (11,164)=35.02, $p<.001$, φc=.14. The first test produced a highly significant p-value of $p<.001$ and a small effect size of φc==.14. This suggests that the results are of high statistical significance but a small level of practical significance. The second test produced a significant p-value of $p<.001$ and a small effect size of φc=.14. This shows that the number of listeners choosing the incorrect options over the correct option was of high statistical significance but little practical significance.

The results for question 2 suggest a substantial number of listeners misinterpreted the sonification by choosing incorrect options and a similarly substantial number of listeners reported that they did not understand the sonification. This indicates that listeners could not effectively link the lowest data point to its time point using the sonification. The results for question 1 and 2 suggest that listeners cannot effectively link data values with time-points when listening to a sonification organised in terms of the design framework.

For question 3 and question 7 listeners were presented with sonifications and asked to determine the half of the year in which the sonifications presented their final data-points. Listeners could select either half of the year or a third choice of "Don't Know". For question 4 and question 8 listeners were presented with sonifications and asked to determine the exact month on which those sonifications ended. Listeners could select any of the 12 months or a 13th option of "Don't Know".

[FIGURE 6]
Figure 6. Data Breakdown for Questions 3 and 7

The data for questions 3 and 7 are presented in figure 6. For question three the correct option was the first half of the year. The results show that listeners had difficulty answering question three with close to equal numbers of listeners selecting the first half of the year option and the second half of the year option. For question seven a clear majority of listeners selected the correct answer, the second half of the year.

Two chi-square goodness-of-fit tests were carried out on the data for question three. The first test included the results for the two halves of the year and the 'don't know' option $\chi 2$ (2,191)=28.34, p<.001, φc=27. The second test considered only the results for the two halves of the year $\chi 2$ (1,167)=.02, $p>.05$, φc=.01. The first test produced a significant p-value of $p<.001$ and a medium effect size of φc=.27. This suggests that the results are of high statistical significance and a medium level of practical significance. The second test produced a non-significant p-value of $p>.05$ and a very small effect size of φc=.01. This suggests that the number of listeners choosing the correct option, the first half of the year, over the incorrect option was statistically non-significant and of almost no practical significance.

Two chi-square goodness-of-fit tests were carried out on the data for question seven. The first test included the results for the two halves of the year and the 'don't know' option $\chi 2$ (1,191)=55.98, $p<.001$, φc=.38. The second test considered only the results for the two halves of the year $\chi 2$ (1,157)=28.59, $p<.001$ φc=.42. The first test produced a significant p-value of $p<.001$ and a large effect size of φc=.38[2]. This suggests that the results are of high statistical significance and a large level of practical significance. The second test produced a significant p-value of $p<.001$ and a medium effect size of φc=.42. This suggests that the number of listeners choosing the correct option, the second half of the year, over the incorrect option was of high statistical significance and medium practical significance.

The results for questions 3 and 7 suggest that a substantial number of listeners can accurately estimate what time-point they are hearing only 50% of the time (once in 2 questions). This result is no better than chance and suggests that listeners cannot effectively identify the half of the year in which a sonification organised in terms of the design framework ends.

[FIGURE 7]
Figure 7. Data Breakdown for Questions 4 and 8

Questions 4 and 8 examined how precisely listeners could pinpoint the exact months on which a sonification ended. The results are presented in figure 7. 52 listeners pinpointed the exact time point correctly for question 4 and 40 listeners pinpointed the exact time point correctly for question 8. For question 8 the number of listeners selecting the correct option and the number of listeners selecting the 'don't know' option were roughly equal.

[2] Effect sizes for Cramer's V are interpreted on different scales determined by df. Hence .38 is a big effect size for 2 df but .42 is only a medium sized effect for 1 df in the results above.

Two chi-square goodness-of-fit tests were carried out on the data for question four. The first test included the results for the twelve-month options and the 'don't know' option $\chi 2$ (12,191)=168.58, p<.001, φc=.27. The second test considered only the results for the twelve-month options $\chi 2$ (12,157)=158.43, p<.001, φc=.3. The first test produced a significant p-value of p<.001 and a medium effect size of φc=.27. This suggests that the results are of high statistical significance and a medium level of practical significance. The second test produced a significant p-value of p<.001 and a large effect size of φc=.3. This suggests that the number of listeners selecting the correct option, March, over the eleven incorrect alternatives was of high statistical significance and large practical significance.

Two chi-square goodness-of-fit tests were carried out on the data for question eight. The first test included the results for the twelve-month options and the 'don't know' option $\chi 2$ (12,191)=135.36, p<.001, φc=.24. The second test considered only the results for the twelve-month options $\chi 2$ (11,149)=95.1, p<.001 φc=.24. The first test produced a significant p-value of p<.001 and a medium effect size of φc=.24. This suggests that the results are of high statistical significance and a medium level of practical significance. The second test produced a significant p-value of p<.001 and a large effect size of φc=.24. This further suggests that the number of listeners selecting the correct option, September, were of high statistical significance and large practical significance.

The results for questions 4 and 8 indicate that a substantial number of listeners can accurately identify the month in which a sonification ends using the framework but an equally substantial number of listeners report that they did not know in which month the sonification ended. This indicates that while the framework is communicatively effective for some listeners, a similar number of listeners do not understand it.

For question 5 and question 9 listeners were presented with sonifications and asked to determine how many time-points had been presented since the beginning of the sonification. Listeners could choose any number between 1 and 12 and a 13th option of "Don't Know".

[FIGURE 8]
Figure 8. Data Breakdown for Questions 5 and 9

Questions 5 and 9 examined how precisely listeners could identify the number of months presented since the beginning of the sonification. The results are presented in figure 8. 51 listeners identified the number of months correctly for question 5 and 47 listeners identified the number of months correctly for question 9. For question 9 the number of listeners selecting the correct option and the number of listeners selecting the 'don't know' option were close in value. Two chi-square goodness-of-fit tests were carried out on the data for question five. The first test included the results for the twelve number options and the 'don't know' option $\chi 2$ (12,191)=160.95, p<.001, φc=.27. The second test considered only the results for the twelve number options $\chi 2$ (11,157)=149.88, p<.001, φc=.29. The first test produced a significant p-value of p<.001 and a medium effect size of φc=.27. This suggests that the results are of high statistical significance and a medium level of practical significance. The second test produced a significant p-value of p<.001 and a large effect size of φc=.29. This suggests that the number of listeners selecting the correct option, three, over the eleven incorrect alternatives was of high statistical significance and large practical significance.

Two chi-square goodness-of-fit tests were carried out on the data for question nine. The first test included the results for the twelve number options and the 'don't know' option $\chi 2$ (12,191)=186.54, p<.001, φc=.29. The second test considered only the results for the twelve-month options $\chi 2$ (11,148)=151.83, p<.001, φc=.3. The first test produced a significant p-value of p<.001 and a large effect size of φc=.29. This suggests that the results are of high statistical significance and a roughly high level of practical significance. The second test produced a significant p-value of p<.001 and a large effect size of φc=.3. This further suggests that the number of listeners selecting the correct option, nine, were of high statistical significance and large practical significance.

The results for questions 5 and 9 indicate that a substantial number of listeners can accurately estimate how many time points have been presented since the beginning of a sonification using the framework. However, a similarly substantial number of listeners report that they did not know how many time points had been presented in a sonification made using the framework. This indicates that while the framework is communicatively effective for some listeners, a similar number of listeners do not understand it.

[FIGURE 9]
Figure 9. Data Breakdown for Questions 6 and 10

Questions 6 and 10 examined how precisely listeners could determine the number of time-points remaining in a sonification of time series data. The results are presented in figure 9. 44 listeners identified the correct number of remaining time points for question 6 and 47 listeners identified the correct number of time points remaining for question 10. For both question 6 and 10, the numbers of listeners selecting the correct options and the number of listeners selecting the 'don't know' options were similar in value.

Two chi-square goodness-of-fit tests were carried out on the data for question six. The first test included the results for the twelve number options and the 'don't know' option $\chi 2$ (12,191)=145.7, p<.001, φc=.25. The second test considered only the results for the twelve number options $\chi 2$ (11,152)=118.47, p<.001, φc=.27. The first test produced a significant p-value of p<.001 and a medium effect size of φc=.25. This suggests that the results are of high statistical significance and a medium level of practical significance. The second test produced a significant p-value of p<.001 and a medium effect size of φc=.27. This further suggests that the number of listeners selecting the correct option, nine, over the eleven incorrect alternatives was of high statistical significance and a medium level practical significance.

Two chi-square goodness-of-fit tests were carried out on the data for question ten. The first test included the results for the twelve number options and the 'don't know' option $\chi 2$ (12,191)=168.58, p<.001, φc=.27. The second test considered only the results for the twelve number options $\chi 2$ (11,148)=158.44, p<.001, φc=.31. The first test produced a significant p-value of p<.001 and a medium effect size of φc=.27. This suggests that the results are of high statistical significance and a medium level of practical significance. The second test produced a significant p-value of p<.001 and a large effect size of φc=.31. This further suggests that the number of listeners selecting the correct option, three, was of high statistical significance and large practical significance.

The results for questions 6 and 10 indicate that a substantial number of listeners can accurately determine the number of time-points remaining in a sonification using the framework. However, a similarly substantial number of listeners report that they did not know how many time-points remained in a sonification made using the framework. This indicates that while the framework is communicatively effective for some listeners, a similar number of listeners do not understand it.

For question 11 and question 12 listeners were presented with sonifications in which the rainfall data for two consecutive months was represented in sequence. They were then asked to determine which month had the highest data value. Listeners had 3 choices, the first month, the second month or "Don't Know".

[FIGURE 10]
Figure 10. Data Breakdown for Questions 11 and 12

The results for questions 11 and 12 are presented in figure 10. 97 listeners selected the correct option, the first month, for question 11. 107 listeners selected the correct option, the second month for question 12. Two chi-square goodness-of-fitness were carried out on the data for question eleven. The first test included the results for the two months the 'don't know' option $\chi 2$ (2,191)=47.41, p<.001, φc=.35. The second test considered only the results for the two halves of the year $\chi 2$ (1,170)=3.38, p>.05, φc=.14. The first test produced a significant p-value of p<.001 and a large effect size of φc=.35. This suggests that the results are of high statistical significance and a large level of practical significance. The second test produced a non-significant p-value of p>.05 and a small effect size of φc=.14. This was of low statistical significance and a small degree of practical significance. This suggests that listeners were confused by the sonification. The high level of statistical significance and large level of practical significance produced by the first test suggest that the number of listeners that did not know which option to choose was of high statistical significance and a large degree of practical significance.

Two chi-square goodness-of-fit tests were carried out on the data for question twelve. The first test included the results for the two months and the 'don't know' option $\chi 2$ (2,191)=55.58, p<.001, φc=.38. The second test considered only the results for the two month options $\chi 2$ (1,168)=.12.6, p<.001, φc=.27. The first test produced a significant p-value of p<.001 and a large effect size of φc=.38. This suggests that the results are of high statistical significance and a large degree of practical significance. The second test produced a significant p-value of p<.001 and a small effect size of φc=.27. This shows that the number of listeners choosing the correct option, the second month, over the incorrect option was of high statistical significance but showed only a small degree of practical significance.
These preliminary results suggest that 50% of the time (once in two questions), a significant number of listeners can effectively estimate the relationship between data-points that are sequentially presented using the design framework. This is not better than chance.

The results for questions 11 and 12 indicate that a substantial number of listeners can effectively estimate the relationship between data-points that are sequentially presented using the design framework 50% of the time (once in two questions). This result is no better than chance and suggests that listeners could not effectively understand the relationship between sequentially presented data-points sonified in terms of the design framework.

5.2 Discussion of Results.

Questions 1 and 2 tested how effectively listeners can link data points to time-points within the design framework. For question 1 a statistically and practically significant number of participants did not know how data values and time-points were related in the stimuli. For question 2 a statistically and practically significant number of listeners misinterpreted the relationship between data values and time-points. These results provide strong evidence that the design framework does not aid listeners in linking data values to time-points. In the conceptual metaphor guide, illustrated in table 3, physical objects were mapped onto data points, changes in physical objects were mapped to changes in data value and vertical path of motion, in the sense of real world vertical movement through physical space, was mapped to the value of the data. These results suggest that in the context of the design framework, these mappings do not reliably hold for the listener. This may have come about because the Doppler shift applied to the data objects interfered perceptually with the listeners' ability to correctly interpret the values of pitch mapped data in cases where the data did not have a high degree of variance.

Questions 3, 4, 7 and 8 tested how effectively listeners could determine what time-point they are hearing when data is sonified using the framework. Questions 3 and 7 tested whether or not listeners could identify the half of the year in which a sonification ended. The results suggest that listeners can accurately identify this only 50% of the time. This does not represent a reliable enough frequency. The results for question seven further suggest that in the cases where the number of listeners selecting the correct option was of high statistical significance and medium practical significance, a similar number of listeners still reported that they did not know how to use the framework. Metaphorical mappings supported by the results for question three are represented in rows 2 to 5 of table 8.

Questions 4 and 8 examined how well listeners could identify the exact months in which data points were presented. The results for these questions suggest that a number of listeners of high statistical significance and a medium degree of practical significance could use the framework to identify the time-point that they were hearing. However, a similarly large number of listeners reported that they did not know how to use the framework.

This suggests that the design framework may be useful in a limited manner for representing a time series in sonification. These results suggest that the design framework is strongest when representing time-points and not as strong when expressing the relationships between data values. In the original *moving times* metaphor, objects are mapped onto times, as illustrated in table 1. In the *moving observer* metaphor locations on the observer's path' are mapped onto times, as illustrated in table 2. In the guide, illustrated in table 3, physical objects were mapped onto 'Data Points'. An analysis of the results for questions 1 and 2 indicates that this proposed mapping may be inaccurate. The analysis of questions 3,4,7 and 8 suggests physical objects might map to times, in the same way that they do for the *moving times* and *moving observer* metaphor[3]. Further, this would suggest that temporary cessation of motion might map to the pausing of the time series, and the speed of motion might map to the speed of the time series' presentation.

Questions 5, 6, 9 and 10 tested how effectively listeners could estimate their temporal distance from the beginning and end points of a time series. Questions 5 and 9 tested distance from the start point while 6 and 10 focused on distance from the end point. Listeners could determine both their distance from start and end points with a strong level of statistical significance and a medium to high degree of practical significance for each of the questions. However similar numbers of listeners also reported that they did not know how to determine their temporal proximity to the start and end points of a sonification using the framework. This suggests that the design framework might be used in a limited manner to give listeners a sense of temporal context in a sonification. The results for questions 9 and 10 are very similar, suggesting that once a listener has determined how many time-points they have already heard, their knowledge of the time series, and/or the design framework, aids in accurately estimating the remaining time-points. Overall these results fall in line with what would be expected based on some of the metaphorical mappings suggested in the conceptual metaphor guide and illustrated in table 7. The accuracy with which listeners judged the distance to sonification end points may suggest that they were performing some basic arithmetic using their prior knowledge of the data and their knowledge of the time-points to which they had just listened. This cannot be said with certainty. Nor can it be said with certainty that listeners were utilising the knowledge that the passing data-objects with the closest proximity to them represented the mid-point of the data. Regardless it can be drawn that for this evaluation, listeners were able to determine their distance from the end-point of a sonification to a limited degree.

Questions 11 and 12 tested how effectively listeners judge the relationships between data points presented sequentially within the design framework. The results suggest that listeners can accurately identify this 50% of the time. This does not represent a reliable enough frequency. These results further support the findings in questions 1 and 2, demonstrating that the design framework is not effective at representing relationships between data values. The results from evaluations 3 to 10 suggest that the framework is better suited to represent points in a time series in a limited manner. This further indicates that a number of the metaphorical mappings suggested in the conceptual metaphor guide,

---

[3] Sonically there may be no difference between sonifications built around either of these two metaphors unless the designer chose to represent the sound of a listener in motion, e.g. walking sounds, or the sounds of vehicular travel, in a sonification built around the *moving observer* metaphor.

where physical objects map to data-points, changes in physical objects map to changes in data value, and vertical path of motion maps to the value of the data, are incorrect.

These results have some implications for the conceptual metaphor guide presented originally in table 3. The results suggest that physical objects map to times rather than data points and that changes in physical objects do not map to changes in data values. The proposed mapping from vertical height to data value is also invalidated by the evaluation. These findings are represented in the improved conceptual metaphor guide presented in table 8. We suggest that this guide may be of use to sonification designers who are working with time series data.
These findings also suggest that the design framework is more suited to representing the passage of points in a time series rather than the value of data points. It offers limited support to listeners incorrectly identifying the time-points with which they are presented, and in estimating their temporal distance from the start and end points of a sonification.

[TABLE 8]
Table 8. Improved Conceptual Metaphor Guide

Some of the results generated in this evaluation might be accounted for in terms of the limits of human memory. There are a number of factors surrounding the listener's abilities in terms of auditory memory, which are not addressed directly by the framework and may account for some of the difficulties listeners encountered during these evaluations. The stimuli used in the evaluations were produced on two time scales. The stimulus for evaluation A consisted of 12 individual unit sequences of 9 seconds in length with a 3-second overlap between the end of one sequence and the beginning of another while the unit sequences for all of the other stimuli were 7 seconds in length with an overlap of 2.33 seconds. With the specifications for evaluation A, the framework can play back audio representing 12 months of data in 1 minute and 16 seconds. The individual unit sequences are too long in duration to be held and processed by echoic memory and each stimulus roughly reaches the capacity of the lower limits of working memory discussed in section 2.2. Similarly, the entire sonification presented in evaluation A is too long to be held and processed as a whole in working memory. These factors may have contributed to the difficulties listeners encountered during the evaluation. Creating faster variants of the framework might help to negate these effects. Such sonifications could present individual stimuli within a 4-second window in order to exploit the meaning-making capacities of echoic memory. They might also group portions of the time series into a task-specific number of chunks that fall within the 10-20 second window of working memory and observe the Miller limit. The way in which the limits of auditory memory are addressed may be different for different sonification designs depending on the task or usage for which the sonification is intended. Furthermore, listeners' attempts to interpret the data from the pitch of the data objects may have been complicated by the pitch shifting used to simulate the Doppler effect. These effects can be accounted for in our revised version of the framework below.
One of the cues that the system offers the listener to help determine their position in the time series (in this case which months data they are hearing) is the relative angle of presentation (in terms of panning) for each data object.
This may be have been a problematic design choice because, as noted previously, MAA is highly dependent on the angle of presentation as well as the frequency content of the sound, especially for higher frequencies (Mills, 1958). Given that the systems employed panning and a data to pitch based mapping strategy the listeners' ability to localise the data objects would have been compromised.
A further consideration here is the question of aural skills. As mentioned in section 2, Miller (1956) points out that listeners can accurately make absolute judgements of the pitches of up to 6 tones but that musically trained listeners with perfect pitch might discern between 50 and 60. Listeners with such abilities may interpret the sonifications described in this article much differently. The framework relies on the listeners' ability to detect precise changes across a number of complex dimensions, from pitch to stereo position, and make comparisons between sequentially presented sounds. While it can be difficult to measure the relationship between aural skills and perception in auditory display (Neuhoff 2011) accounting for differences between musical and non-musical users will be important to the further development of this framework.

5.3 Recommendations and Reflections

Given these results, we have a number of recommendations for representing time series data. Some of these recommendations have been codified here in the form of an improved design framework. This improved design framework represents an iteration upon the original informed by the results of the evaluations. Further iterative cycles of design and evaluation are required to advance the development of this design framework. The updated framework only makes prescriptions for the representation of temporal context cues, or 'time markers', as such listener no longer needs to make determinations about the value of the data from sounds that are also intended to represent time-points. The results show that the design framework did not aid listeners in linking data values to time-points or in representing relationships between non-time series data values. As such the improved framework is solely concerned with representing the time series data, allowing the designer to deal with the associated data values in a manner of their own choosing. This is achieved by denoting each individual time-point in the data with sonic symbols of the designer's choice (e.g., pitch or timbre objects) to create sonic time markers. These time markers replace the data objects in the original framework. Data is not mapped to the time markers they are simply used as context cues to denote the passage of unit sequences. Over the course of a single unit sequence, a time marker will seem to pass by the listener, moving from a position in front of the listener to a position behind them. This trajectory, from front to back, represents the passage of a specific unit sequence as illustrated in figure 11. These passing time markers are organised from left to right, relative to the listener's position, across the horizontal axis of the stereo field. The time marker for the earliest point in the time series appears to the extreme distant left of the listener. The midpoint of the time series appears at the listener's location and the last point in the time series appears to the extreme distant right of the listener. This configuration, represented in figure 12, creates a sequence of time-indexed sonic vectors. These time markers make one entire pass by the listener's position along the y-axis (front to back) over the course of a single time-point and one entire pass by the listener's position along the z-axis (left to right) across the entire time series.

[Figure 11]
Figure 11. Time Markers

The observed data values relating to the time points can then be presented via pitch-mapped sonification in the centre of the stereo field. This decouples the time series representation form its corresponding data value representation and allows the framework to be superimposed onto a sonification to lend a sense of temporal context. Accounting for the limitations of echoic memory a maximum unit sequence of 4 seconds should be used. Given this limited unit sequence length, it makes sense to represent no more than two values with any single data object, one before the pitch shift and one after. This should help to ensure that values are clearly represented to a listener. The specifics of how exactly this achieved is a question for the designer. For example in a pitch-mapped PMSon, these data points could either be presented as static pitches, or a smooth interpolated sequence of frequencies between both data points. Similarly overlaps become redundant in this improved framework, as each time mark is presented independent of the data sonification and as such only a single time point in the data can be represented at any given time in the sonification. We also suggest adhering to Miller's 7 ± 2 rule as an upper limit on the number of data objects in any one sonification, this ensures that the design stays within the limits of working memory. Lastly, a single sonification of this type should fall within the 10-20 second window outside of which working memory decays.

[Figure 12]
Figure 12. Time Series Spatialisation

To highlight the operation of this improved framework, consider a simple pitch-mapped PMSon of the rainfall data presented previously. The improved framework can be applied here to impose a grid of time markers over the simple pitch-mapped sonification. Time markers should be represented with sonic symbols that have a distinct timbre which differentiates them from the main sonified data stream, and also makes them easier to spatialise. In this case they are generated using fonction d'onde formantique (fof) synthesis techniques to create a fuller spectrum that is distinct from the sine tones used in the PMSon. The original PMSon will be presented at an angle of 0° in the stereo field. The time markers are spatialised in the same manner as the data objects were in the previous iteration of the framework. As the data for the first month is played a time marker will be presented on the listeners far left at an angle of -90°. Consecutive time markers will sound in a sequence of x° increments (in this case 15°), reaching the listener's position at 0° and continuing on to +90°. Reverb is applied to the time markers in the same manner it was applied to the data objects in the previous framework. Reverb is modulated such that time markers sound as though they approach the listener from the left and retreat on the right. The Doppler effect is also applied to the time markers in the same manner as it was applied to data objects in the original framework. Each individual time marker begins in the area in front of the listener and passes to the area behind them over the course of a single time unit (in this case a month). That being said the problem of front/back perception in the stereo field is tricky and while the Doppler effect may provide an additional cue to help in addressing this, it certainly does not solve the problem definitively. You can hear an example of the application of this iteration of the framework in the third supplementary audio file.

The sonifications presented in this article involved the mapping of data values to pitch in a simple sine wave. The decision was made to use sine waves for the research presented in this paper for a number of reasons. The first reason is that this approach allowed us to maintain tight control over the number of variables in our evaluation design and to make concrete statements about experimental findings that may not have been possible if we had to account for the idiosyncrasies involved in the use of more complex sound sources. For example, using recorded sound can introduce spatial information that would interfere with attempts to spatialise the audio. Likewise, certain sounds, e.g. those with more uniform spectral content, lend themselves better to spatialisation than others (Middlebrooks & Green 1991).
It was also hoped that this approach would lead to the production of more genearlisable results that would not be inextricably linked to the specific particularities of the sounds chosen, as we imagined may become the case when dealing with more complex and idiosyncratic sounds. Making sure that this approach was accessible to sonification designers and did not require advanced sound design knowledge and skills was also a consideration. Another reason for this approach is that the framework was designed with conceptual metaphor theory in mind and according to Lakoff and Johnson (1987) conceptual metaphors make sense of abstract and unfamiliar domains of experience by relating

them, at an unconscious level, to familiar domains of embodied experience. Sine tones represent an abstract and unfamiliar sonic domain for the average listener, and the literature suggests that, given the spatial cues applied, listeners are likely to interpret these somewhat unfamiliar sounds in terms of movement through physical spaces. Furthermore, it was initially felt that using familiar sounds to represent the data might be closer to the process of cross-domain mapping also discussed in the embodied cognition literature (Lakoff and Johnson 1999). In cross-domain mapping content from one domain of embodied human experience is associated content with from another but this process does not involve the critically important reframing process that takes place in conceptual metaphor (Lakoff and Johnson 1999). Regardless, of the exact cognitive factors involved the inclusion of more complex and familiar domains of sound in future iterations of the framework is likely to enhance comprehensibility. Auditory display research on ecological psychoacoustics (Walker and Kramer, 2004) and embodied cognition (Roddy and Furlong, 2014) has suggested more alternative design strategies to pitch mapping which may be more comprehensible to a listener. Because a listener rarely encounters a sine wave during their everyday life, they may be somewhat unfamiliar with them and as such may have difficulty interpreting a sonification or auditory that uses them. Furthermore, because the auditory system we have today evolved over millions of years in environments populated with complex sonic signals, we are better adapted to interpreting and deriving meaning from these more complex signals. For example, as noted previously, sounds recorded from, or synthesised to simulate, natural environments tend to have a broader frequency spectrum and flatter range of amplitude values which make them easier to localise for a listener. We therefore intend to explore the use of sounds that are familiar to the user from their everyday environments and activities over the abstract sine tones employed here in future iterations of the framework. We suggest that auditory display designers consider these rich sonic domains in their work also.

6.1 Limitations & Future Work

One limitation of the design framework is that it can only handle a shorter time series. The effectiveness of the framework is constrained by two factors. The first is the resolution of the human auditory perceptual and cognitive system in localising sounds. The second is the designers' ability to represent a graded sense of distance to that auditory perceptual and cognitive system. The applications of the framework in this article focused on the rainfall data measured in mm at Dublin airport over the course of 2014 as represented in table 4. There is a relatively short span of only 12 time-points represented in this dataset. For datasets with a larger number of time-points listeners will not be able to discern small incremental changes in the spatial cues used. This can be addressed in a limited manner by treating ranges of sequential time-points as the sub-units of a finer timescale within the framework's time units. These sub-units are represented by tn(.01) to tn(.99) in figure 12. These subunits can be configured to suit the designer's needs. This solution is still limited by the resolution of the perceptual and cognitive auditory systems discussed in section 2.4. It can only be effective up to the threshold of MAA where listeners can still distinguish one time-point from the previous time-point. As such the framework is better suited to providing a topological gist of the time series to a listener.

We also suspect that it might be most useful when both the unit sequence and the sonification as a whole are allowed to represent standard units of time. For example, representing data spanning a week in which unit sequences of 1 day or data spanning a month with a unit sequence of 1 week. As pointed out in the previous section, in order to ensure that the data is as intelligible to the listener as possible it may also be preferable to represent no more than 2 values in any one unit sequence: one value before the Doppler shift and one value after the shift.

One possible application we envision for a future iteration of the framework is in the reporting of weekly metrics where data is sampled at intervals of twice a day. Data of this type can be represented consistently with our improved framework specification. IoT network deployments which support environmental monitoring (for flooding, noise levels, energy consumption etc.), generally receive small packets of data at regular intervals from smart measurement devices. The design framework could be applied when reporting weekly changes in these networks. This would involve representing two daily readings for a single device over the course of a single week. Dividing a sonification of this data using 7 time markers might be manageable for the listener. Initial readings could sound before the Doppler pass and the second reading could sound after, giving the listener an overview of each day's data. Sonifications of this type might be useful for getting a quick overview of a week's worth of data from a specific device. This approach could be used to represent the growth in the number of devices registered on a network over the course of a week. Similar use cases outside of the IoT context might include weekly reports tracking changes in stock prices, changes in weather data and sales or manufacturing numbers for a business.

The design of the initial framework presented here was driven by a top-down, theory led approach. The aim was to create a prototype framework informed by embodied cognition principles and to undertake some initial testing that might determine the viability of this approach. Evaluation participants were recruited from a crowdsourcing platform, and this, coupled with our evaluation design, made it difficult to elicit an in-depth understanding of what did and didn't work for them. The next steps for this project will involve further design iterations and user evaluations pursed via a participatory design (PD) methodology that will directly integrate users into the design process (Sanders 1992 ; Schuler and Namioka 1993). This approach will invole a larger number of participants than the evaluation presented here and will allow us to gain a deeper understanding of the aspects of the system that work, those that do not, and, critically, why that may be the case. It will also help to keep the future iterations of the system grounded in the users' needs and are adapted to the users' abilities.

Future work might also explore spatialisation using the ambisonic B-format to develop an interactive application that will allow the user to zoom in and listen to shorter sections of the time in greater detail or zoom out to get an overall gist of larger time spans. Allowing for different levels of analysis in this way might allow the framework to be scaled up for application to larger and more complex datasets.

6.2 Conclusions

This article has drawn some principles from research on conceptual metaphors to inform a design framework for the sonification of time series data. The framework is intended to provide a sense of temporal context in time series sonification. It draws from two unconscious systems of conceptual metaphor discussed in the embodied cognition literature (Lakoff and Johnson, 1980; Cox, 1999; Johnson and Larson, 2003;) whereby our understanding of time is framed in terms of motion through physical space. Some implications about possible strategies for time series data sonification are drawn from these metaphors and represented in our conceptual metaphor guide. We then suggest a design framework for time series data sonification on this basis. The framework is evaluated with an empirical listener testing methodology. Analysis of the results suggests that the initial framework specification provides the listener with a limited sense of temporal context. It aids the listener in estimating which time-point they are currently hearing, how many time-points they have already heard, and how many time-points they have left to hear. Recommendations are then made for an improved conceptual metaphor guide and an improved framework, which accounts for and builds upon the findings of the evaluations.

References

Antle, A. N., Corness, G., Bakker, S., Droumeva, M., van den Hoven, E., & Bevans, A. (2009). Designing to support reasoned imagination through embodied metaphor. In *Proceedings of the seventh ACM conference on Creativity and cognition* (pp. 275-284). ACM.

Antle, A. N., Corness, G., & Bevans, A. (2011). Springboard: Designing Image Schema Based Embodied Interaction for an Abstract Domain. *Whole Body Interaction Human-Computer Interaction Series,* 7-18. doi:10.1007/978-0-85729-433-3_2

Benesty, J. (2017). *Springer handbook of speech processing*. Berlin; Heidelberg: Springer.

Bødker, S. (2015). Third-wave HCI, 10 years later---participation and sharing. *Interactions, 22*(5), 24-31. doi:10.1145/2804405

Bødker, S., & Klokmose, C. N. (2016). Dynamics, Multiplicity and Conceptual Blends in HCI. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI 16*. doi:10.1145/2858036.2858530

Brazil, E. & Fernström, M. (2006). Investigating concurrent auditory icon recognition. *Proceedings of the 12th International Conference on Auditory Display (ICAD 2006)*, London, UK, 20-23 June.

Cabrera, D., & Tilley, S. (2003). Parameters for auditory display of height and size. Georgia Institute of Technology. *Proceedings of the 9th International Conference on Auditory Display (ICAD 2003),* Boston, MA, 6-9 July.

Clark, T. (1987). Echoic Memory Explored And Applied. *Journal of Services Marketing, 1*(2), 41-48. doi:10.1108/eb024707

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104*(2), 163-191. doi:10.1037//0033-2909.104.2.163

Cowan, N. (1995). *Attention and memory an integrated framework*. New York: Oxford University Press ; Oxford.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87-114. doi:10.1017/s0140525x01003922

Cox, A. (1999). The metaphoric logic of musical motion and space (Unpublished Doctoral dissertation). University of Oregon.

Darwin, C. J., Turvey, M. T., & Crowder, R. G. (1972). An auditory analogue of the sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology, 3*(2), 255-267. doi:10.1016/0010-0285(72)90007-2

Diniz, N., Coussement, P., Deweppe, A., Demey, M., & Leman, M. (2012). An embodied music cognition approach to multilevel interactive sonification. *Journal on Multimodal User Interfaces, 5*(3-4), 211-219. doi:10.1007/s12193-011-0084-2

Dourish, P. (2004). *Where the action is the foundations of embodied interaction*. Cambridge, MA: Bradford Books.

Fairfax, T., Laing, C., & Vickers, P. (2014). Network Situational Awareness: Sonification and Visualization. In M. M. Cruz-Cunha, I. M. Portela, J. Mull & A.A. Escamilla (Eds.), *Handbook of research on digital crime, cyberspace security, and information assurance*. Hershey, PA: Information Science Reference. (p. 334)

Fitch, T. W., & Kramer, G. (n.d.). Sonifying the body electric: Superiority of an auditory over a visual display in a complex, multi-variate system. In G. Kramer (Ed.), *Auditory display: Sonification, audification, and auditory interfaces* (pp. 307-326). Reading, MA: Addison-Wesley.
Flowers, J. H. (2005). Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions. *Proceedings of the 11th International Conference on Auditory Display (ICAD2005), Limerick, Ireland, July 6-9*.

Gayer, W. W. (1989). The SonicFinder: An Interface that Uses Auditory Icons. *ACM SIGCHI Bulletin, 21*(1), 124. doi:10.1145/67880.1046601

Gerzon, M. A. (1992). *The design of distance panpots: Presented at the 92nd convention 1992 March 24-27 Vienna*. New York, N.Y: Audio Engineering Society.

Grond, F., & Berger, J. (n.d.). Parameter mapping sonification. In T. Hermann, A. Hunt & J.G. Neuhoff (Eds.), *The sonification handbook* (pp. 363-397). Berlin: Logos Verlag.

Hermann, T. (2002). *Sonification for exploratory data analysis* (Unpublished Doctoral thesis). Bielefeld, University. Retrieved from: http://sonification.de/publications/media/Hermann2002-SFE.pdf

Hurtienne, J. (2009). *Image schemas and design for intuitive use: Exploring new guidance for user interface design* (Unpublished Doctoral thesis). Berlin, Technische Universität.

Imaz, M., & Benyon, D. (2007). *Designing with blends: Conceptual foundations of human-computer interaction and software engineering*. Cambridge, MA: MIT Press.

Johnson, M. L., & Larson, S. (2003). "Something in the Way She Moves"-Metaphors of Musical Motion. *Metaphor and Symbol, 18*(2), 63-84. doi:10.1207/s15327868ms1802_1

Kolarik, A. J., Moore, B. C., Zahorik, P., Cirstea, S., & Pardhan, S. (2015). Auditory distance perception in humans: A review of cues, development, neuronal bases, and effects of sensory loss. *Attention, Perception, & Psychophysics, 78*(2), 373-395. doi:10.3758/s13414-015-1015-1

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York, NY: Basic Books.

Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York, NY: Basic Books.

Macaranas, A., Antle, A. N., & Riecke, B. E. (2012). Bridging the gap: attribute and spatial metaphors for tangible interface design. *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction - TEI 12*. doi:10.1145/2148131.2148166

Middlebrooks, J. (1991). Sound Localization By Human Listeners. *Annual Review of Psychology, 42*(1), 135-159. doi:10.1146/annurev.psych.42.1.135

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 101*(2), 343-352. doi:10.1037/0033-295x.101.2.343

Mills, A. W. (1958). On the Minimum Audible Angle. *The Journal of the Acoustical Society of America, 30*(4), 237-246. doi:10.1121/1.1909553

Neuhoff, J. G. (2011). Perception, Cognition and Action inAuditory Displays. In T. Hermann, A. Hunt, & J. G. Neuhoff (Eds.), *The sonification handbook*. Berlin: Logos Verlag.

Oh, J., & Wang, G. (2012). Evaluating crowdsourcing through amazon mechanical turk as a technique for conducting music perception experiments. *Proceedings of the 12th International Conference on Music Perception and Cognition*.

Olson, H. F. (1967). *Music, physics and engineering: 2D. ed*. New York: Dover.

Perrott, D. R., & Saberi, K. (1990). Minimum audible angle thresholds for sources varying in both elevation and azimuth. *The Journal of the Acoustical Society of America, 87*(4), 1728-1731. doi:10.1121/1.399421

Razavi, B., Ehsani, P., Kamrava, M., Pekelis, L., Nangia, V., Chafe, C., & Parvizi, J. (2015). The brain stethoscope: A device that turns brain activity into sound. *Epilepsy & Behavior, 46*, 53-54. doi:10.1016/j.yebeh.2015.02.051

Rimland, J., Ballora, M., & Shumaker, W. (2013). Beyond visualization of big data: A multi-stage data exploration approach using visualization, sonification, and storification. *Next-Generation Analyst*. doi:10.1117/12.2016019

Roddy, S., & Furlong, D. (2014). Embodied Aesthetics in Auditory Display. *Organised Sound, 19*(1), 70-77. doi:10.1017/s1355771813000423

Roddy, S., & Bridges, B. (2018). New Directions in Third Wave Human-Computer Interaction: Volume 2 - Methodologies. In M. Filimowicz & V. Tzankova (Eds.), *New Directions in Third Wave Human-Computer Interaction: Volume 2 - Methodologies* (Vol. 2). Cham: Springer International Publishing.

Roddy S., & Bridges B. (2018b) Addressing the Mapping Problem in Sonic Information Design through Embodied Image Schemata, Conceptual Metaphors and Conceptual Blending. The Journal of Sonic Studies. 17(1). ISSN 2212-6252

Rohrhuber, J. (2010). S–Introducing sonification variables. In *Proceedings of the Supercollider Symposium*.

Rusconi, E., Kwan, B., Giordano, B., Umilta, C., & Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition, 99*(2), 113-129. doi:10.1016/j.cognition.2005.01.004

Sanders, E. B. (2010). Converging Perspectives: Product Development Research for the 1990s. *Design Management Journal (Former Series), 3*(4), 49-54. doi:10.1111/j.1948-7169.1992.tb00604.x

Schuler, D., & Namioka, A. (2009). *Participatory design: Principles and practices*. Boca Raton: CRC.

Vickers, Paul (2012) *Ways of listening and modes of being: electroacoustic auditory display.* Journal of Sonic Studies, 2 (1). ISSN 2212-6252.

Walker, B. N., & Kramer, G. (2004). Ecolological Psychoacoustics and Auditory Displays. *Ecological Psychoacoustics,* 149-174. doi:10.1016/b978-012515851-0/50007-2

Walker, B. N., & Kramer, G. (2005). Mappings and metaphors in auditory displays. *ACM Transactions on Applied Perception, 2*(4), 407-412. doi:10.1145/1101530.1101534

Walker, B. N. & Nees, M. A. (2011). Theory of sonification. In T. Hermann, A. Hunt & J.G. Neuhoff (Eds.), *The sonification handbook* (pp. 9-39). Berlin: Logos Verlag.

Waterworth, J., & Riva, G. (2014). *Feeling Present in the Physical World and in Computer-Mediated Environments*. Palgrave Macmillan.

Wolf, K., & Fiebrink, R. (2013). SonNet: A Code Interface for Sonifying Computer Network Data. *Proceedings of the 13th Conference on New Interfaces for Musical Expression*, (pp. 503-506).

Woodworth, R. S.(1938). *Experimental Psychology*. New York: Holt, Rinehart & Winston.

Worrall,D.(2015). Realtime Sonification and Visualisation of Network Metadata (The NetSon Project). *Proceedings of The 21st International Conference on Auditory Display,* (pp. 337-339).

Figure Legend

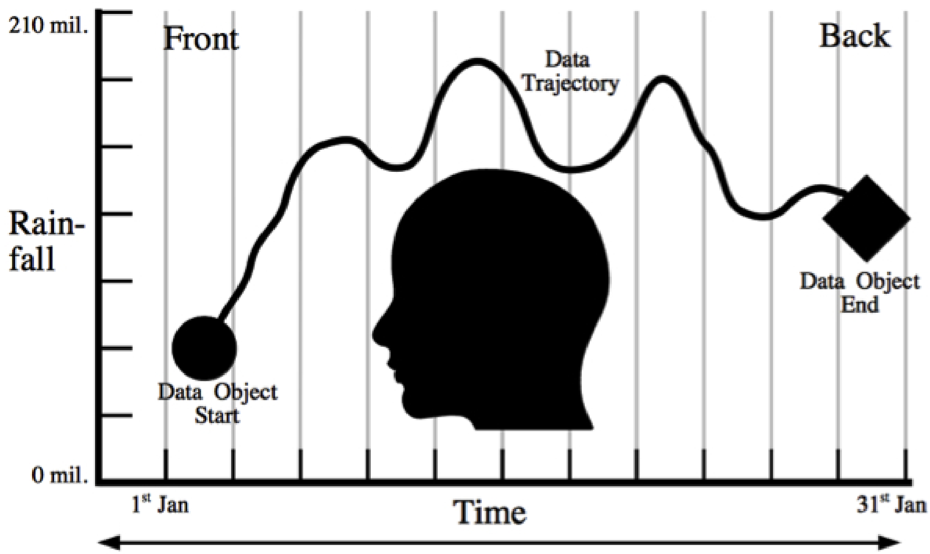


Figure 1. Unit Sequence from Front to Back

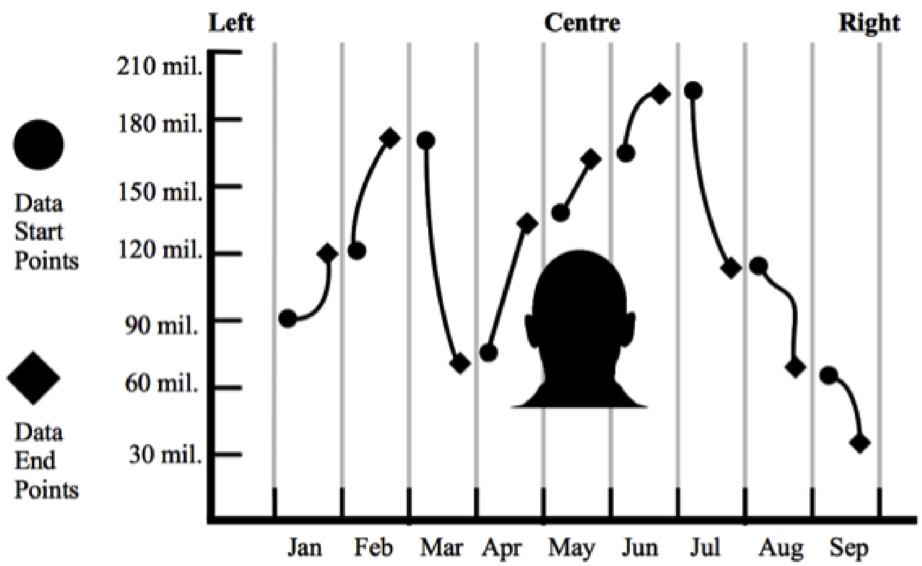
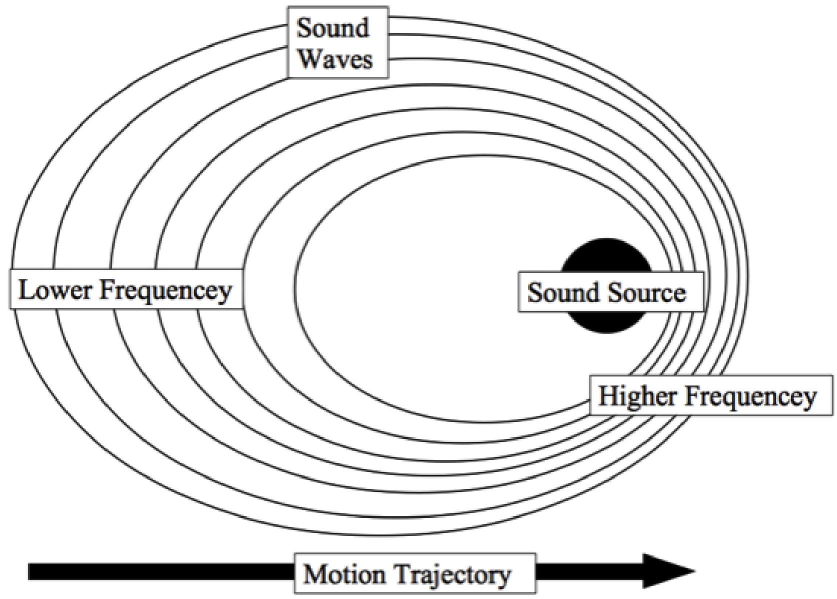
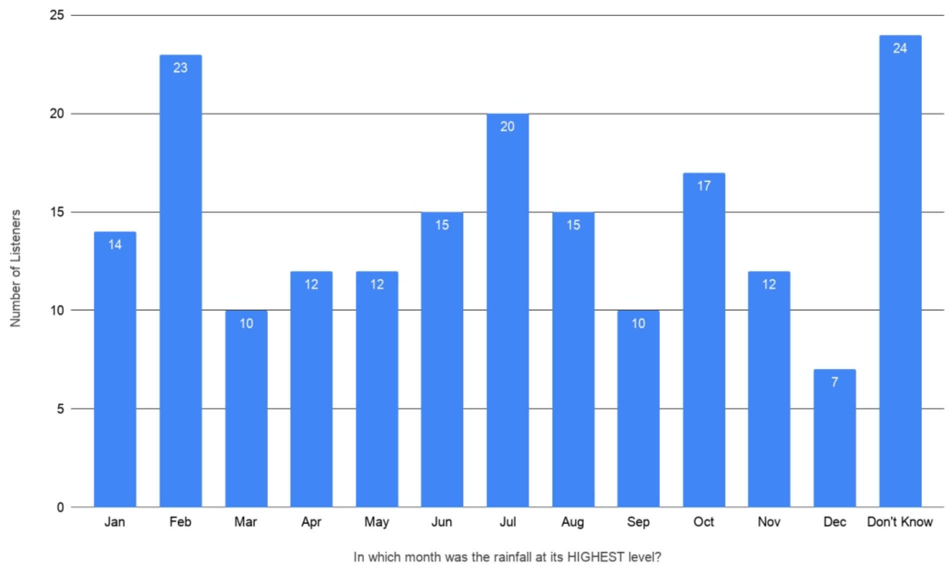


Figure 2. Time Series from Left to Right



Doppler Effect



Caption : Figure 4. Data Breakdown for Question 1

541x334mm (72 x 72 DPI)

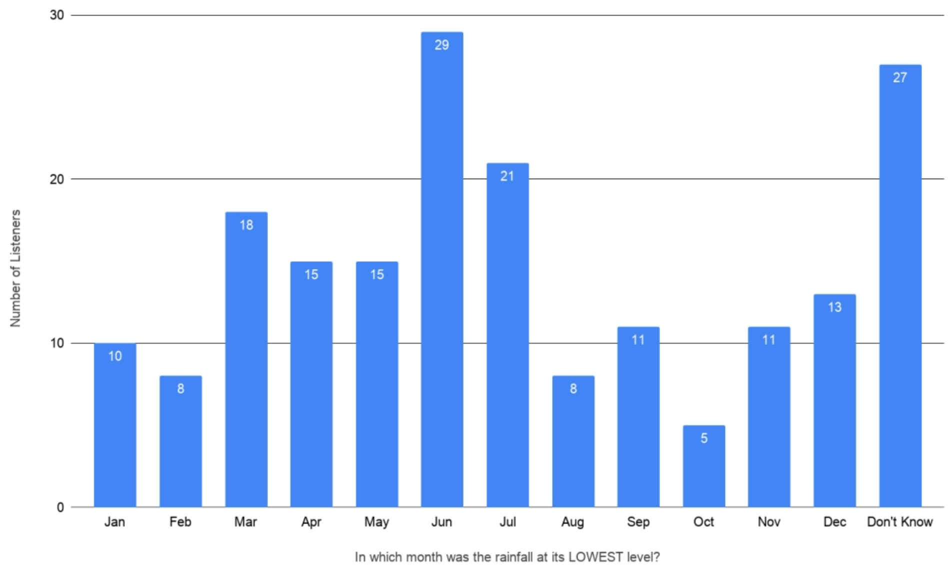


Figure 5. Data Breakdown for Question 2

541x334mm (72 x 72 DPI)

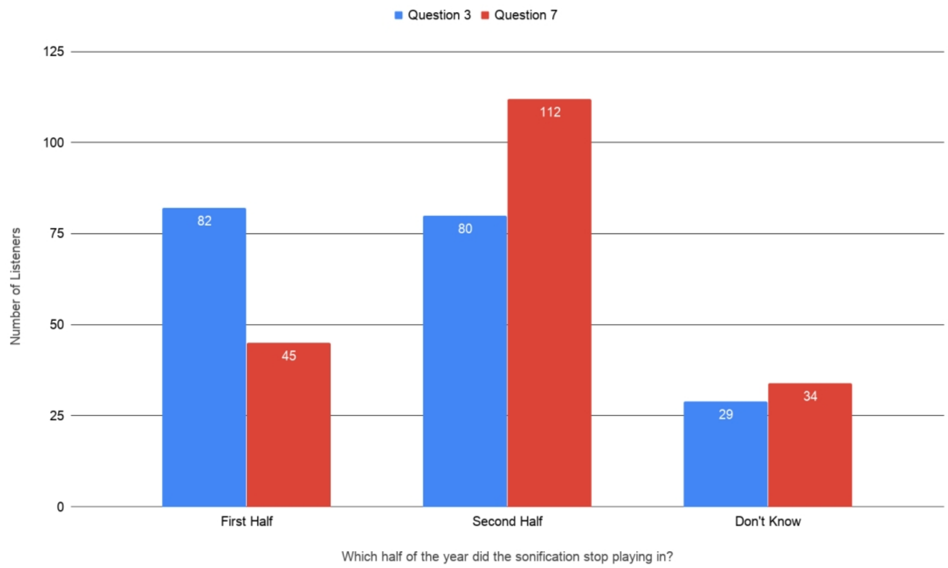


Figure 6. Data Breakdown for Questions 3 and 7

541x334mm (72 x 72 DPI)

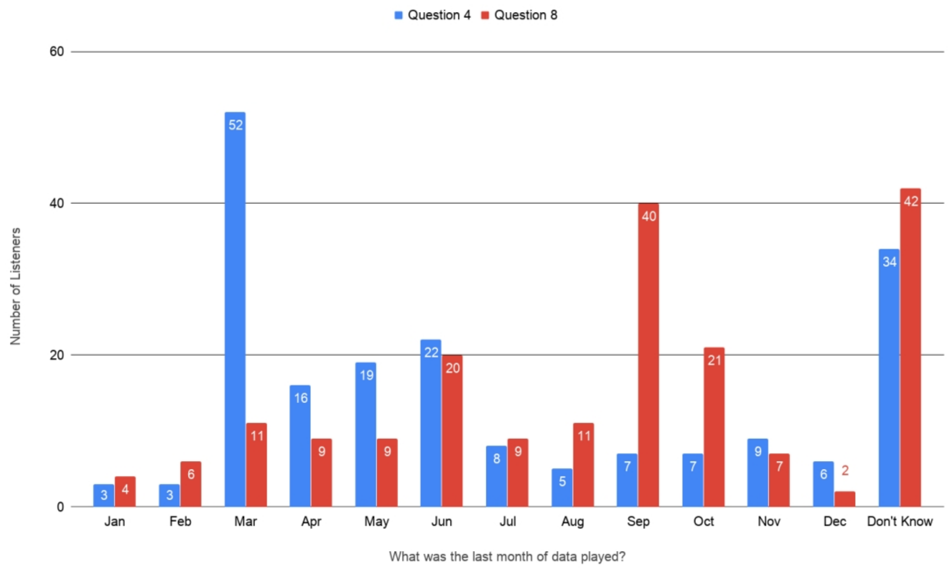


Figure 7. Data Breakdown for Questions 4 and 8

541x334mm (72 x 72 DPI)

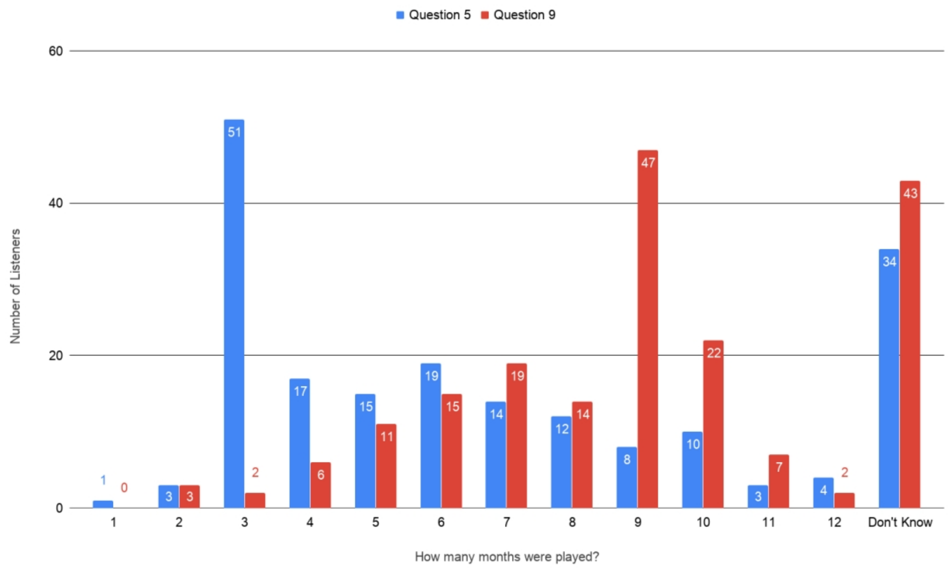


Figure 8. Data Breakdown for Questions 5 and 9

553x341mm (72 x 72 DPI)

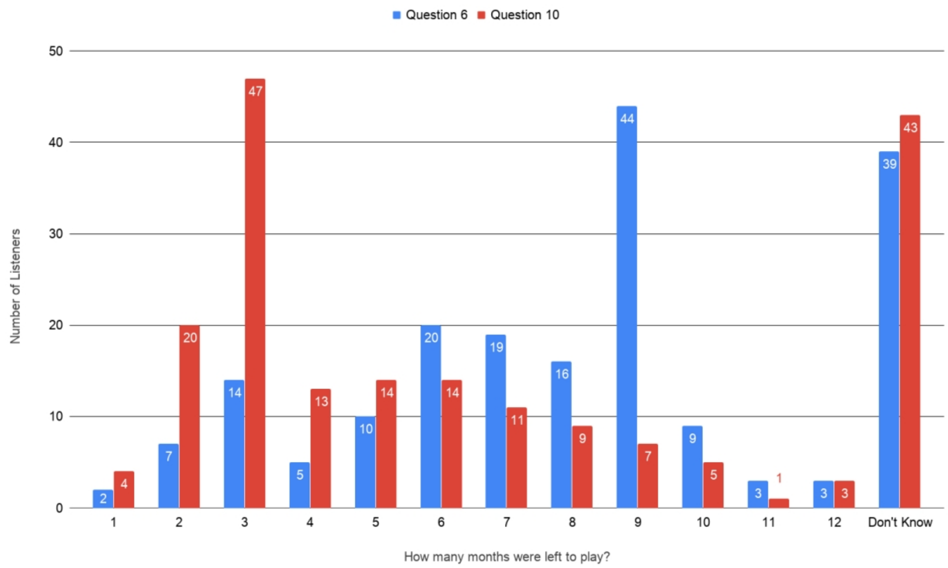


Figure 9. Data Breakdown for Questions 6 and 10

553x341mm (72 x 72 DPI)

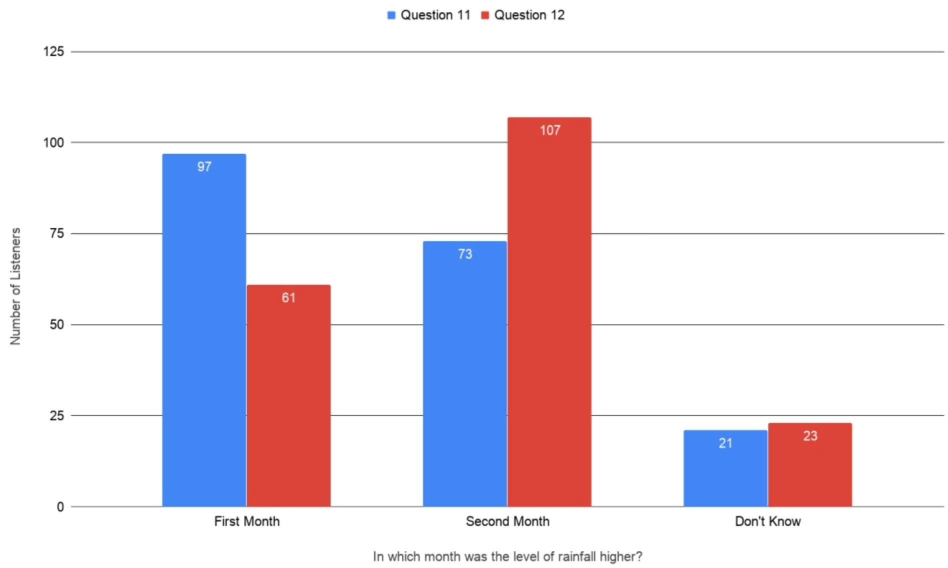


Figure 10. Data Breakdown for Questions 11 and 12

541x334mm (72 x 72 DPI)

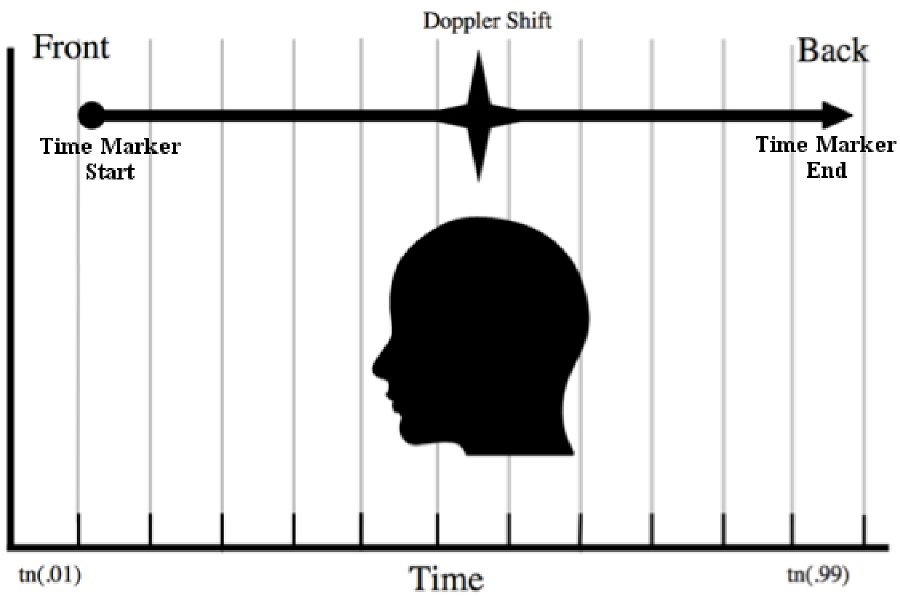


Figure 11. Timer Markers

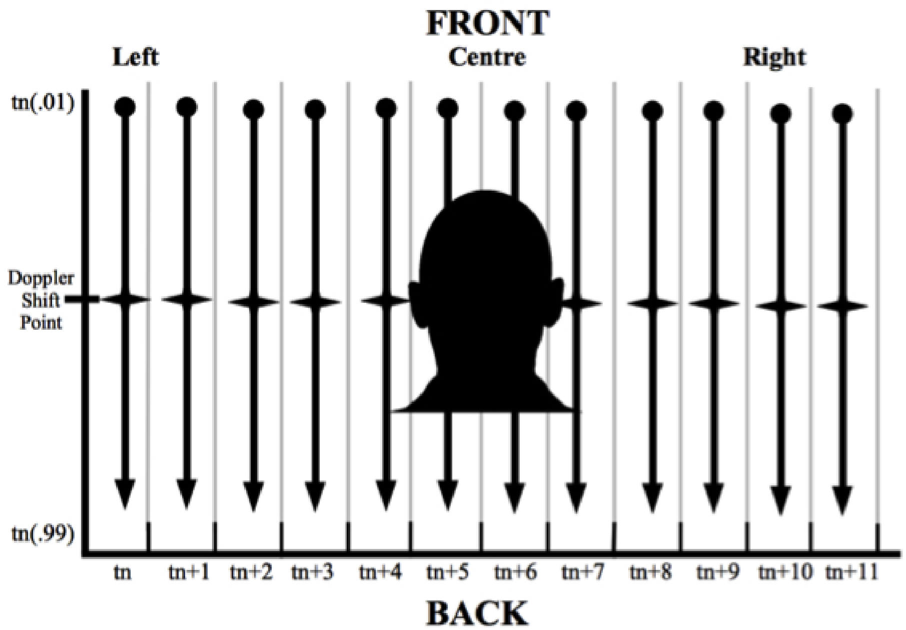


Figure 12. Time-Series Spatialisation

Source Domain (Space)		**Target Domain (Time)**
Objects	=	Times
Motion of objects past the observer	=	The passage of time
Location of the observer	=	The present
Space in front of the observer	=	The future
Space behind the observer	=	The past

Table 1. The Moving Times Metaphor

Source Domain (Space)		**Target Domain (Time)**
Locations on observer's path	=	Times
Motion of the observer	=	The passage of time
Distance moved by the observer	=	Amount of time passed"
Location of the observer	=	The present
Space in front of the observer	=	The future
Space behind the observer	=	The past

Table 2. The Moving Observer Metaphor

Data	
Data Point	>>
Change in data value	>>
Value of the data	>>
Passage of single unit of time	>>
Passage of entire time series	>>
Midpoint of time series	>>
Start of time unit	>>
End of time unit	>>
Start of time series	>>
End of time series	>>
Pausing of time-series	>>
Speed of time-series presentation.	>>

Table 3. Conceptual Metaphor Guide

Metaphor: Physical Motion	
Physical object	>>
Change in physical object	>>
Vertical path of motion	>>
Movement past listener Front/Back	>>
Movement past listener Left/Right	>>
Listener’s horizontal position	>>
Listener's front	>>
Behind listener	>>
Listener's left	>>
Listenrer's right	>>
Temporary cessation of motion	>>
Speed of Motion	>>

Sound

Sound ▪
Change in sound ▪
Pitch of sound
Front/Back movement of sound ▪
Left/Right movement of sound ▪
Listener's location ▪
Percieved area in front of listener ▪
Percieved area behind listener
Percieved distant left of listener ▪
Percieved distant right of listener ▪
Pausing of sonification
Speed of sonification

Month	Jan	Feb	Mar	Apr	May	Jun
Rainfall	101.6	88.5	53.7	34.2	91.5	36.2

Table 4. Rainfall in millimetres recorded at Dublin Airport over 2014

Jul	Aug	Sep	Oct	Nov	Dec
35	173	26.5	90.2	140.9	55.9

Month	Pan Degrees	Reverb Wet: Dry Ratio	Reverb Pre-delay Time in Milliseconds	Early Reflection Amplitude
Jan	-90	0:100	100	-30
Feb	-75	20:80	80	-24
Mar	-60	40:60	60	-18
Apr	-45	60:40	40	-12
May	-30	80:20	20	-6
Jun	-15	!00:00	0	0
Jul	15	!00:00	0	0
Aug	30	80:20	20	6
Sep	45	60:40	40	12
Oct	60	40:60	60	18
Nov	75	20:80	80	24
Dec	90	0:100	100	30

Table 5. Control Parameters for Left to Right Distribution

	Answers + Don't Know	Answers Only
Q1	$\chi2$ (12, 191)=21.42, p<.05, ϕc=.1	$\chi2$ (11, 167)=15.87, p>.05, ϕc=.09
Q2	$\chi2$ (12, 191)=43.74, p<.001, ϕc=.14	$\chi2$ (11,164)=35.02, p<.001, ϕc=.14
Q3	$\chi2$ (2,191)=28.34, p<.001, ϕc=27	$\chi2$ (1,167)=.02, p>.05, ϕc=.01
Q4	$\chi2$ (12,191)=168.58, p<.001, ϕc=.27	$\chi2$ (12,157)=158.43, p<.001, ϕc=.3
Q5	$\chi2$ (12,191)=160.95, p<.001, ϕc=.27	$\chi2$ (11,157)=149.88, p<.001, ϕc=.29
Q6	$\chi2$ (12,191)=145.7, p<.001, ϕc=.25	$\chi2$ (11,152)=118.47, p<.001, ϕc=.27
Q7	$\chi2$ (1,191)=55.98, p<.001, ϕc=.38	$\chi2$ (1,157)=28.59, p<.001 ϕc=.42
Q8	$\chi2$ (12,191)=135.36, p<.001, ϕc=.24	$\chi2$ (11,149)=95.1, p<.001 ϕc=.24
Q9	$\chi2$ (12,191)=186.54, p<.001, ϕc=.29	$\chi2$ (11,148)=151.83, p<.001, ϕc=.3
Q10	$\chi2$ (12,191)=168.58, p<.001, ϕc=.27	$\chi2$ (11,148)=158.44, p<.001, ϕc=.31
Q11	$\chi2$ (2,191)=47.41, p<.001, ϕc=.35	$\chi2$ (1,170)=3.38, p>.05, ϕc=.14
Q12	$\chi2$ (2,191)=55.58, p<.001, ϕc=.38	$\chi2$ (1,168)=.12.6, p<.001, ϕc=.27

Table 6. Analyses of Results

Source Domain (Physical Motion)		**Target Domain (Sonification)**
Movement past listener Front/Back	=	Passage of unit of time
Movement past listener Left/Right	=	Passage through data set
Listener's horizontal position	=	Midpoint of the data
Start/end of motion Left/Right	=	Beginning/end of time series
Start/end point of motion Front/Back	=	Beginning/end of time unit

Table 7. Mappings Supported by the Findings for Question 3

Data

Time Points

- Passage of single unit of time
- Passage of entire time series
- Midpoint of time series
- Start of time unit
- End of time unit
- Start of time series
- End of time series

Pausing of time-series

Speed of time-series presentation.

Table 8. Improved Conceptual Metaphor Guide

▪
▪
▪
▪
▪

>>
>>
>>
>>
>>
>>
>>
>>
>>
>>

Metaphor: Physical Motion

Physical object	>>
Movement past listener Front/Back	>>
Movement past listener Left/Right	>>
Listener’s horizontal position	>>
Listener's front	>>
Behind listener	>>
Listener's left	>>
Listenrer's right	>>
Temporary cessation of motion	>>
Speed of motion	>>

Sound
Time Markers
Front/Back movement of time marker
Left/Right movement of time marker
Listener's location
Percieved area in front of listener
Percieved area behind listener
Percieved distant left of listener
Percieved distant right of listener
Pausing of sonification
Speed of sonification