Issues in Auditory Display.

A.C.Eldridge
email: alicee@sussex.ac.uk
Creative Systems Lab., School of Informatics,
University of Sussex, Brighton, BN2 9QH, UK
Tel: +44 1273 877645
Fax: +44 671320

January 14, 2005

Abstract

Auditory displays have been successfully developed to assist data visualization in many areas, but have as yet received little attention in the field of Artificial Life. This paper presents an overview of existing design approaches to auditory display and highlights some of the immediate challenges for the field. Examples from recent experiments are used to illustrate the importance of considering factors such as data characteristics, data-display mappings, perceptual interactions within and between display modalities, and user experience and training in designing new visualization tools. It is concluded that whilst further research is needed to develop generic design principles for auditory display, this should not stand in the way of exploration of bespoke designs for specific applications.

Keywords: auditory display, sonification, visualization tools, complex systems, cellular automata, training.

1 Introduction.

Graphical representations enable us to spot patterns or trends in data by removing the clutter of absolute numerical values and exposing important interval relations [4]. Existing theoretical and experimental evidence suggests that in certain situations, it may be easier to comprehend salient structures using an auditory rather than a visual representation and that sound can be a beneficial addition to traditional graphical displays.

The use of sound in human-computer interfaces is not new (eg [45], [48]), simple alarm signals to focus attention or announce the completion of a task have been used for many years. Tools such as the Geiger counter demonstrate that sound can be more effective than a visual display for enabling comprehension of time-varying structural details [55]. The superiority of auditory representations for some tasks can be understood in terms of the perceptual characteristics that are peculiar to the auditory domain (see section 2.1). These theoretical possibilities are supported by recent experimental investigations of practical applications of auditory display for a range of analysis and monitoring tasks (see section 2.2).

This paper explores the potential for auditory display as an analysis tool in Artificial Life (ALife) research. In the first section, an overview of existing research in the field of auditory display is given. General advantages of auditory displays are described in terms of perceptual characteristics of the auditory system, and results from existing research are used to demonstrate some of the advantages of the approach. In section 3, the most common approaches to developing auditory displays are outlined. Some of the immediate challenges facing the field are considered and are illustrated by examples from recent experiments using example visualization problems relevant to ALife research. It is concluded that although a considerable amount of further research is required before auditory display emerges as a fully theoretically informed field, immediate benefits can be gained through individual exploration.

2 Why Listening Works.

2.1 Characteristics of Auditory Perception.

One of the principle advantages of auditory perception, is that unlike visual perception, there is no requirement for specific user orientation either physically, visually or of attentional focus. This makes auditory displays ideal for monitoring tasks and for scanning, both for specific predicted patterns, or unforeseen anomalies [31]. The natural alerting properties of sound also suggest application for audio in multimodal displays, for instance, in focusing attention on key areas of complex visual displays.

Besides extending the possible display area beyond the visual focus, the nature of acoustic perception means that certain data types may be intrinsically easier to comprehend through the acoustic channel. It has been suggested that multidimensional data in general [6], [19] and logarithmic or time-varying data in particular [6], may be more effectively presented to the ear than to the eye. Speech-based evidence of selective-attention (eg [22]) suggests that the auditory system may be capable of monitoring data structures embedded in other more static signals which would be too noisy to apprehend visually (see section 2.2.2). The superior temporal resolution of the acoustic system [49] suggests that fast changing or transient events that may be blurred or entirely missed visually could be detected in the simplest of auditory displays. Sensitivity to temporal characteristics also enables discrimination between periodic and aperiodic events and the propensity to detect salient patterns, even when subject to radical transformation, highlights the potential for using sound in pattern recognition tasks or data mining.

The auditory system also possesses a remarkable ability to detect interaural time differences, perceived subjectively as changes in the apparent location of sound source. Advances in creating individualized head-related transfer functions (HRTFs) [63] facilitate the exploitation of this ability in sonification design. More generally, the auditory systems' capacity to *compare* two streams of data presented binaurally could be capitalized upon, for example in scanning paired data sets for correlations.

2.2 Advantages of Auditory Display.

These characteristics confer advantages on auditory display, both in terms of efficacy in apprehending certain types of data, and efficiency in terms of user effort or time.

2.2.1 Decreasing Subjective Workload.

In applied settings such as medicine and engineering, the limiting factor in analysing and monitoring large data-sets is no longer computer performance, but the constraints imposed by human perception. In particular, long procedures and analyses induce fatigue which typically manifests itself as attention deficits, sensory motor habituation and distorted perception of time flow [28]. Researchers at the University of Belgrade developed a sonic representation of electroencephalograph (EEG) recordings by controlling left-right pan using the left-right brain hemisphere EEG power symmetry. In a pilot test in clinical settings, neurophysiologists are reported to have gained better insights into global brain electrical activity, and reductions in mental fatigue compared to those without the sonification support.

Similar results from Human Computer Interaction (HCI) research support the proposition that the addition of audio displays can reduce subjective workload. Brewster et al [10] investigated the application for sound to enhance usability of mobile Personal Digital Assistant (PDA) devices. Participants were required to maximize profit by trading shares using either a simple auditory display, or a line graph to monitor price changes. Performance was equal in both modes, but participants reported a significant decrease in workload with the sonification, presumably because this meant the visual display could be used solely for trading.

2.2.2 Facilitating Comprehension of Certain Types of Data.

As well as reducing workload, presentation of data in audio rather than graphical formats can promote comprehension of certain structures. An ongoing project at Loughborough University is examining the use of 'auralisation' of code as a debugging technique. The CAITLIN system has been developed which produces auditory output by mapping certain Pascal language structures (IF, CASE, WHILE, REPEAT, FOR) to musical motifs. Early research showed that users were able to describe the structure of simple programs from the auralisation alone [56], and subsequent research suggests that in certain situations auralisations assist in locating bugs for novice programmers [56], [57].

Accounts from significant scientific events also corroborate theoretical evidence that the auditory system is capable of discerning data structures that are impervious to graphical analysis. During the Voyager 2 space mission, the cause of the problems encountered as the craft approached Saturn could not be established, graphical depictions of incoming data from on-board meters showing pure noise. But when a synthesiser was used to transform the data into sound, a 'machine gun' effect was heard at the critical period, which led to the realisation that the craft was being bombarded with electromagnetically charged micrometeoroids [31]. The discovery of the 'quantum whistle' has also been attributed to the use of auditory display. The oscillations predicted by quantum theory could not be detected using a visual oscilloscope, however, transformation of the data into an acoustic signal created a faint whistle, providing the first experimental support for theoretical predictions [46].

Of key interest in the current context is the potential for audio representations in facilitating comprehension of complex dynamic systems. The stethoscope is a standard medical instrument, with which medical students learn to *listen* to blood pumping through veins and gases bubbling in the intestines. Experimental results show that in a simulated operation, medical students provided with eight dynamic variables describing the health of a patient presented in audio, out-performed those given visual, and even audio-visual displays [17]. Results from other medical and engineering investigations into auditory display support the idea that cycles, rhythms, patterns and short events are particularly amenable to acoustic analysis

[38].

Existing psychoacoustic evidence is corroborated by conspicuous success stories and as the relevant audio technologies become increasing available, auditory display is developing from a field of enquiry into one of application. An established annual conference held by the International Community of Auditory Display (http://www.icad.org) hosts discussion of design approaches and applications for auditory display in a range of disciplines. Beside medical [17], [37] and cockpit alarm applications, auditory interfaces in assistive technologies for the visually impaired [36], [29] are proving particularly useful; the advantages of sonic enhancements in mobile computing [8], VR and desktop computing interfaces have been recognised, and there is an increasing interest in the use of auditory display for scientific visualization, for example in the analysis of engineering simulations, [38] or comprehension of seismic data [23], [12].

3 Approaches to Sonification.

> "Sonification is . . . the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation"
> [33] p.3.

At the inaugural conference of auditory display, attendees were invited to design an auditory display to support the categorization of six dimensional data sets, as either dirt or gold. The most notable results from this exercise was the enormous variation in display design. The variety of approaches seen at the conference is characteristic of the field, and is due in part to the range of applications. Approaches have been classified along a spectrum between analogic and symbolic [31] p.21-29. Analogic representations are characterized by a one-to-one mapping from data to display dimensions, creating an intrinsic correspondence between structures in the data and the representation medium as in the Geiger counter. In contrast, symbolic representations do not necessarily preserve the structural relationships of the data being represented, but categorically denote some aspect of it, exemplified by the auditory alarm. The most common methods of auditory display are outlined below with reference to their suitability as methods for scientific visualization.

Earcons are arbitrary acoustic motifs that can be used in combination to convey more complex messages. In the simplest form, these are used as alarm signals, such as the BIOS 'beep' announcing the state of the hardware, but syntactic combinations can be developed to signify a more complex repertoire of events. For example a short sine wave at 440Hz may symbolize 'file A', and another set of pitches designated to states of that file such that the note 'A' followed by an 'E' denotes that 'file A is closed'.

Auditory Icons, like earcons were developed to provide feedback from the Graphical User Interface (GUI) but rather than using arbitrary motifs, selections are modeled on 'every day' sounds, for example task progress may be represented by the sound of a bottle being filled with liquid [21]. This is approach is seen to aid learnability by evoking common metaphors, although there is no experimental support for this [35].

Icons and earcons deployed in mobile computing device menus have been shown to have positive effects in shortening performance times and decreasing errors [9]. The inclusion of a 'beep' to signal the end of an evolutionary run, or other salient point in a simulation may similarly save time, but the categorical nature of both icons and earcons limits their use as analysis tools for complex data.

Audification describes the direct translation of data into the auditory domain [31]. This guarantees an intrinsic correspondence between structures in the data and the representation medium such as is found in the Geiger counter. This technique has proved particularly effective in the analysis of seismic data, in which the frequency of vibrations is accelerated to the range of human hearing [23], [12]. The approach has obvious potential for time series analysis, being particularly meaningful if the data originates from a dynamically evolving system, such as the network described in section 4.1.1.

Model-based sonification is a relatively new approach being developed principally at Bielefeld University [25], [26], [24]. Taking an ecological stance [20], the approach aims to capitalize upon the auditory system's ability to extract information from the acoustic signal created by our interaction with the world (we can tell instantly how hard or rough a surface is by hitting or scraping it). A model-based sonification therefore requires the reframing of the data as a virtual scenario and the definition of a virtual physics. Once defined, these laws can be used to measure how the elements react to external excitation by the user. Information about the data can then be derived from these measurements. For example, data points could be conceived as planets and a gravitational force defined. Particles could then be introduced into the data space to probe the gravitational potential at various points, from which the structure of the data set as a whole could be inferred. This approach has proved successful for several data pre-processing tasks such as analyzing clusters in vectorial data, and exploring the separability of a vectorial data set prior to a classification task [25].

Parameter mapping is the most common method used for scientific data analysis and involves the mapping of data dimensions onto display dimensions. Display dimensions may be parameters which determine the tone (timbre) of the sound (such as modulation depth or speed, spectral evolution, attack time etc.), parameters such as frequency, volume or tempo, or even high level musical structures such as motifs.

Developments in synthesis techniques and physical modeling offer potential for more abstract mappings and manipulation of perceived spatial position is also proving successful. Some of the major advantages of this approach are that multivariate representations are possible, escaping the three-dimensional restraints of graphical displays, and that sonifications can be rapidly developed using existing software tools which enable exploration of many potential data-display mappings.

4 Issues in Auditory Display Design.

The diversity of design approaches reported at the first ICAD meeting was paralleled by a huge variation in the effectiveness of the auditory displays produced: in subsequent tests, users ability to identify key features of the data set ranged from chance to highly significant. As Bly points out [6], the effectiveness of any display will be influenced by the design of the display itself but also the characteristics the data being analyzed. We suggest that the perceptual abilities of the user, and their familiarity with the display type are also important contributory factors. In this section, the design and testing of simple auditory displays for two contrasting dynamic systems is outlined to illustrate the importance of aspects such as characteristics of the data being visualized, data to display mappings, display modality, user experience and learning.

4.1 Example Sonifications of Complex Adaptive Systems

Auditory displays for two different types of systems were developed. One is a ten-node homeostatic network which produces continuous-time real valued outputs. The other is a two-dimensional Cellular Automata, which is updated synchronously, producing discrete, binary data.

Using a traditional graphical display, comprehension of the overall state of the homeostatic network at any point is difficult as it demands the perceptual integration of many transient outputs. A simple 'audification' was developed by translating the outputs of each node into pitch variations. Users were asked simply to comment on how easily they could perceive the overall state of the system using the auditory display, in comparison with graphical representations.

For the CA, the typical graphical representation of cell states *is* an effective means of monitoring the global system state, the focus here then, was on the influence of musical experience on the relative effectiveness of the visual, auditory and combined audio-visual (AV) representations of the CA. A more abstract auditory representation was developed by transforming the spatial patterns into temporal (rhythmic) patterns, and converting certain statistical properties of the production rules into pitch variations. Users were presented with audio, visual and combined AV displays, and asked to classify the qualitative state of the system.

4.1.1 System One: Homeostatic Network

Fig. 1 about here

The network consists of multiple interconnected units, where the output of each unit is determined by the weighted outputs of all connected units in the previous iteration (see Fig.1, left). The system behaves either chaotically, with all outputs varying wildly, or settles to a stable state where all outputs converge to a point or limit cycle. These limit cycles vary in length, such that for large systems it is hard to distinguish visually between limit cycles and random fluctuations, either when plotted on line graphs or mapped to dynamic sliders on a GUI. Increasing the size of the network magnifies this problem making it hard to determine if, and when the system stabilizes. In addition, for larger networks, the time taken to stabilize increases [1], creating large data sets that can be time-consuming to visualize. The fact that we can monitor multiple audio streams simultaneously, suggested that the qualitative state of the system may be more easily perceived from an acoustic representation.

Sonification Scheme. The auditory display is based upon the translation of the real valued outputs of individual nodes in the network into continuous pitch variations. The central position of each unit is assigned a starting pitch at discrete intervals (Fig. 1). Each output is then scaled to produced continuous frequency deviations from this central pitch, creating dynamic microtonal harmonies (Fig. 2). Limit cycles are therefore heard as repetitions of specific sets of frequencies, or loops, that are readily differentiated from the random series of pitches produced by chaotic states[1].

Fig.2 about here...

Advantages. In a series of informal demonstrations and presentations, users agreed that the auditory representation facilitated the comprehension of system dynamics. This was especially true for networks comprising 6-20

[1] Audio files and an example application can be downloaded at: http://www.informatics.sussex.ac.uk/users/alicee/Sonification.html

units (smaller networks could be easily monitored using dynamic components on a GUI, for larger units, the pitch range became too extreme to be comfortable). Users expressed surprise at the clarity of 10 unit system. This may be due to habituation to continuous tones [11] emanating from static units, which 'thins down' the harmonies, enabling the listener to focus on the changing variables.

One of the noted advantages of the auditory presentation, particularly for larger networks, was that it reduced the workload compared to dynamic graphical displays, which were tiring to focus on for any length of time, or even static line graphs which required concerted scanning to analyze. This in turn meant that the addition of the audio, enabled the system to be run at higher speeds at which the dynamic graphic display alone was uncomfortable to monitor or simply incomprehensible.

The alternative mode of presentation also seemed to promote novel insights into the system for both the designer and subsequent users. The audio display enabled rapid exploration of system dynamics for a range of parameter settings, revealing a mathematical error in the code which produced erroneous behavior that had been missed on previous graphical debugging. Further, in a demonstration to Cognitive Science MSc students, several commented that playing with the system provided an understanding of the nature of homeostatic systems that had previously evaded them.

4.1.2 System Two: Cellular Automata

Fig. 3 here

1D CA can be categorized into one of three classes: complex, chaotic or ordered[2]. Although these classes can be distinguished visually according to the patterns produced by a 2D representation of the binary outputs, the pattern detection capabilities of the auditory system suggest that these states may also be effectively presented in an auditory display.

Sonification scheme. The sonification scheme used employs two sets of mappings. One transforms the familiar spatial patterns of the CA into temporal patterns, creating distinct types of rhythms for each class (Fig 3). The other converts simple statistical properties of the rule look-up table into pitch values, creating dynamic harmonies. This is illustrated in Fig. 4. Full details can be found in [15].

Fig 4. here (ish)

[2]These are equivalent to Wolfram's four classes [62] where classes one and two are both ordered.

Perceptual Experience and Practice. One of the main aims of this investigation was to examine the effect of musical experience on the effectiveness of different display modalities. In a controlled experiment, two groups of students, musicians and non-musicians, completed a categorization task using audio-only, visual-only and AV displays. The main result, shown in Fig. 5, showed a differential effect of display type on performance according to musical experience. Musicians performed best in the AV condition, non-musicians in the visual only condition. Scores for both groups were lowest in the audio only condition.

Fig. 5 here

Data collected during the trials also highlights the importance of training for tasks involving novel displays. Although not explicitly manipulated, the number of practice examples each participant viewed before starting the trials was recorded. Overall scores for both groups show a significant correlation with number of practice examples viewed.

Whilst no firm conclusions can be drawn from this work, these results highlight many issues central to auditory display, and suggest some important areas for further research. There are two immediate questions that arise. Firstly, why the auditory display for the homeostatic network was apparently more effective than the graphical, whereas for the CA, it was the less effective. Secondly, why musical experience affects preferred display modality in the CA classification task. Further, more rigorous investigations are needed, but it seems likely that the differences lie in the nature of the systems, the mappings used, and experience-based perceptual differences in the users.

4.2 Data Characteristics.

Both tasks reported here involve some form of pattern recognition, but there are some important differences in the temporal characteristics of the systems and their dimensionality. For the homeostat, appreciation of the state of the network demands recognition of patterns across continuous time and transient outputs in multiple dimensions. In the CA, however, the current state of the system is most easily appreciated when successive states accumulate, and the recent history of system can be seen at a glance in a two dimensional array. Given the characteristics of auditory perception noted in section 2.1, it is perhaps no surprise that the outputs of the homeostatic network were easier to appreciate in the auditory representation. In contrast, the transient nature of sound means that the immediate history of the CA system cannot be presented as easily in audio as it can in a two dimensional graphical display.

The importance of considering the structure of the data under examination has been highlighted previously by Hayward [23]. He suggested that audification of seismic data is more successful than that of other data such as stock market figures, due to a shared physics: "a seismic recording will sound like recording of natural environmental sounds, because sounds transmitted through air (acoustic waves) have similar physics to seismic vibrations transmitted through the earth (elastic waves). The direct, physically consistent, playback can take advantage of human experience with natural sounds" p.93 [23].

Whilst direct audification may more readily achieve more natural results, there seems no a priori reason that other data types may not be successfully sonified. Perhaps the crucial element, as Barrass and Kramer point out is that "relevant changes in the data should ensure a change in what is perceived. Changes in what is perceived should signify meaningful changes in the data." p.25 [3].

4.3 Data-display Mappings.

The key to accurate perceptions of the data, is the development of intuitive and unambiguous mappings from data to display parameters. The audification of temporal data such as seismic vibrations, or network outputs can be simply compressed, producing a simple one-to-one mapping that preserves the inherent data structures. In the examples described here, the time-varying outputs from the homeostat were transformed directly into microtonal pitch variations, giving a one-to-one mappings from data to display dimension. In the CA, however, the mappings were more complex, although based on data relations, the transformations may have been too abstract to be immediately comprehensible. A more effective means of representing the evolution of the patterns in the CA may be to map each element in the array to a pitch value, and increase the speed at which the data is presented in audio, such that patterns in the data are perceived as timbral, rather than rhythmic and melodic variation. The periodic patterns arising from ordered rules, would produce a more harmonic tone, chaotic patterns producing a more noise-like signal. Increasing the speed would also reduce the problems arising from the transience of the acoustic signal. Such a mapping would preserve the inherent synchronous updates and go some way in overcoming the lack of persistence of sound.

This problem is typical for situations in which we wish to build a sonification model for data that is not intrinsically time-based and therefore demands a more abstract or complex set of mappings. Successful designs demand consideration of the psychological meaningfulness of the resulting signal. Currently, most mappings reflect subjective preference, at best evoking common metaphor - such as increases in frequency with temperature - in an attempt to produce mappings that are compelling [33]. Such metaphors are limited however and the mapping procedure for most variables is far from intuitive [58]. For example, should physical size be represented by an increase or decrease in either pitch or loudness of a sound? Such differences in specific data-sound mappings have been shown to affect reaction time and accuracy in monitoring tasks (ibid). However even for common physical dimensions, there seems to be little consensus over preference for particular mappings or their direction [59]. Further research is needed in order to develop sets of design principles to help create mappings that are psychologically meaningful.

4.4 Perceptual Interactions Within Display Dimensions.

Even when intuitive mappings are developed, the limited number of orthogonal dimensions in sound space potentially create perceptual interactions which can distort the way relations within the data are perceived. Numerous studies have demonstrated that the auditory dimensions of pitch, loudness and timbre interact perceptually (eg [40]). Even within one dimension, there appear to be perceptual asymmetries for rising and falling intensities of equal magnitude, e.g. subjects report larger absolute changes in volume when it is increasing, rather than decreasing in level [42]. Recent research has shown that these same interactions and asymmetries occur even when mapped onto data dimensions [43]. Values of stock prices and trading volumes were mapped onto pitch and intensity of an audio signal, and participants were instructed to make judgments of relative changes in trading figures according to perceived changes in the sounds. When both auditory dimensions changed in the same direction, perceived variation in the target variable was reported to be greater than for incongruent changes of the same magnitude.

Timbral parameters are similarly susceptible to interaction, such that linear changes can have unpredictable non-linear perceptual effects. For example, our perception of the 'brightness' of a sound is determined by several factors including the attack time, and spectral evolution. This means that a bivariate display, in which one variable is mapped to the position of the spectral peak and another to the attack time of a static harmonic tone will not be heard as a simple 2D space, as many different combinations of these two variables can create a perceptually equivalent level of brightness. Indeed it has been suggested that a true balanced multivariate parameter mapping may not be possible in practice [32].

Although these interactions may cause problems if data is mapped to continuous parameters, the use of discrete timbral variations can be effective [18]. Using contrasting acoustic textures, much like employing different colors in a graphical display, increases the number of dimensions that can be represented by high level audio dimensions and if carefully designed can prevent masking effects, allowing attention to be equally divided.

Interactions could potentially be employed for positive effect. Being able to present multiple streams creates the possibility for exploiting gestalt grouping principles such as auditory streaming [7]. Users of the auditory display for the homeostatic network described here, reported that the auditory display was perceptually less ‘complex’ than the equivalent graphical representation. This may be due to the fact that co-varying nodes were perceived as a unified stream. Potential distortions of data relations due to the nonlinearity and limited availability of orthogonal acoustic dimensions could be minimized by further research. Awareness of potential interactions could then be considered at the design stage, and certain effects could even conceivably be exploited to highlight salient data trends.

4.5 Cross-Modal Interactions in Multi-Modal Displays

As well as offering an alternative to traditional graphical displays, the unique characteristics of each sense could be combined in multimodal displays. From an ecological perspective, multimodal systems may ‘be more efficient because they better represent real life and the complexity of real life experiences’ [34]. However, debate over the advantages of multimodal presentations in comprehension, learning and memory is long and unresolved: there are situations in which sound can enhance visual displays by providing an extra channel of communication [60], [61], yet cross-modal interactions can cause interference as well as synergetic effects and multimodal displays are not always the most effective [55].

The most commonly reported interference effects arise from situations were contradictory information is presented to each sense. Perhaps most famous is the McGurk effect [39], where perception of a speech phoneme is altered by dubbing it onto a video of a speaker saying a different phoneme. More recently, conflicting audio-visual cues have been shown to create perceptual bias [51], illusions [53] and even cross-modal after effects [30].

Development of **complementary** designs arguably offer most potential by harnessing the distinct properties of sight and hearing for synergetic effects, for example in increasing information capacity in the case of restricted display size [8], or providing detail [41], extra dimensions, or focusing attention in complex displays [50], [37][3], see also [2].

Claims over the benefits of **redundancy** in multimodal displays, principally in education literature, have also been made on the basis that multiple encoding, or cue-summation improves retention, recall, and understanding of contents [16], [13], [52], although this remains a contentious issue. A crucial area of research is in ascertaining under what circumstances cross-modal interactions occur, and which factors lead to synergetic or interference effects.

[3]soundbytes can be found at
http://www.icad.org/websiteV2.0/Conferences/ICAD96/proc96/kramer.htm

4.6 The Effect of Musical Experience on Auditory Perception.

One factor that potentially impacts upon the efficacy on any display with an acoustic element, but has received little attention, is musical experience. Physiological and psychological differences between musicians and non-musicians have been demonstrated [47], and differences in EEG dimensionality between classical and popular music listeners point to the psychophysiological nature of this difference [5]. Yet although frequency change is one of the most widely used dimensions in auditory display, and pitch perception is amongst the most widely researched topics in audition, it is only recently that the effect of musical expertise in simple perceptual, as well as conceptual judgments of pitch has been illustrated. In a controlled experiment, Neuhoff and Wayand [44] tested participants of varying levels of musical experience, and found that musicians reported significantly greater pitch changes than non-musicians for the same interval. In addition errors in judgments of *direction* of frequency change, were significantly greater for non-musicians.

These findings have obvious implications for the development and application of auditory displays, yet have received surprisingly little attention. Results from the trials reported here, suggests that the effectiveness of different displays types was influenced by musical experience. In the CA classification task, musicians' scores were highest when given a combined AV display. This fits with research suggesting that redundant or complimentary representations facilitate comprehension. However, for the non-musicians, the addition of the auditory information seemed to make the task more difficult. Further analysis of the results for this group (see [15]) showed that non-musicians experienced particular difficulty interpreting the audio display. So whilst the auditory and visual elements in the AV display apparently reinforced each other for musicians, perceptual differences may have rendered the AV cues contradictory for the non-musicians.

The differential effect of presentation mode according to musical experience, seems to reflect experience-based perceptual differences. It seems likely that similar individual differences exist in the population at large. It is important to establish the extent and nature of these differences so that individual differences can be considered in developing suitable mapping and scaling procedures.

4.7 The Importance of Training for Novel Displays.

One pertinent line of approach would be to consider the role of learning in acoustic perception. The benefits of learning in enhancing the efficiency of auditory displays have been recognized in areas such as sonar operation [27], and assistive technologies for blind users [14], and have been noted in the context of scientific visualization [33]. The significant correlations between practice time and scores observed in the CA trials reported here support these findings. The effect of practice and familiarity with novel displays must be considered in the assessment and application of auditory display as a scientific visualization tool.

5 Summary.

Toward a Theoretically Informed Approach to Auditory Display Design. Research in auditory perception reveals characteristics of the auditory system that suggest that certain data structures may be readily comprehended via the acoustic channel. Practical application and experimental investigation in a variety of fields corroborate these possibilities, and promote consideration of the potential for auditory display in the real time visualization of complex, dynamic systems, which remains problematic under graphical approaches.

There now exist a substantial number of individual successes in the field, but auditory display is a nascent research area, and still lacks a theoretical basis for many aspects of display design. The test cases described here highlights some important issues that deserve further consideration. Successful application of auditory display inevitably demands careful consideration of whether it can best supplant or supplement visual presentations. Informed decisions can be made on the basis of existing psychophysical research demonstrating the unique specialties of each perceptual sense, but further research is needed to identify which data features are perceptually most salient to each sense, and how to apply this knowledge to display design.

The development of generic principles of display design comparable to those for existing visualization schemes [54], demands further research in a number of areas. Ongoing research in auditory perception, particularly in auditory scene analysis, auditory memory, and the role of attention in extracting information from sound, as well as further research in multimodal perception are essential in forming predictive principles to guide effective representation of the relevant data structures. Results reported here highlight the importance of ascertaining the extent of individual differences in auditory perception, and the effects of practice in increasing the efficacy of novel display types. The issues must be investigated before crucial research into mapping and scaling functions can be sensibly continued.

Although a substantial amount of research is needed before universally predictive guidelines for display design can be drawn up, as the examples given have shown, the benefits of auditory display can be reaped through careful task-led design. The sonified homeostatic network described in section 4.1.1, illustrates the use of straightforward sonification of high dimensional time-varying data. Whilst precise quantitative analysis of system behavior may still ultimately require numerical analysis, and detailed inspection of graphical output, auditory representation can facilitate the debugging, initial comprehension and exploration of a system. The unique affordances of separate sense may be beneficially combined in specific analysis tasks, but the introduction of an alternative mode of presentation increases the choice of tools at all stages of research.

The promotion of ad hoc trial-and error applications may seem ill advised in the absence of predictive guidelines, but many of the problems that hinder the development of generic design principles are inconsequential to the individual user, for whom subjective preference in design choices offer the best solution. The ubiquity of multimedia desktop computers, and availability of sound editors and synthesis packages, makes the practical implementation of an auditory display a trivial task. Indeed the development of the field as a whole may benefit from such personal explorations, as the accumulation of individual successes are vital for stimulating widespread acceptance, funding, support and research time necessary to maximize the potential of auditory display.

Conclusions. Psychoacoustic evidence, and the success of existing applications in other fields, all suggest that certain types of systems, may be easier to understand by listening to, rather than looking at the output data. Possibilities of perceptual interactions, individual differences in perceptual capabilities, and the absence of many psychophysical details all impact on the development of universal theoretical guidelines. However, awareness of potentials and pitfalls and the ease of sound-production in current desktop computing, promote individual exploration of personalized auditory displays, offering alternative solutions to analysis tasks at all stages of research.

6 Resources.

The following sites offer examples, relevant research and tools for sonification:

The International Community of Auditory Display is a forum for discussion and presentation of widespread applications of auditory display, including scientific visualization. Most conference proceedings from the last 10 years are available online at http://www.icad.org

ACM Sigsound hosts links to a many sound-related resources such as DSP, auditory research, courses of study and various sound-related homepages.
http://www.acm.org/sigsound/sigsoundlinks.html

Relevant fundamental psychoacoustic research is available at the Acoustic Society of America site : http://asa.aip.org

Music programming languages
CSound http://www.csounds.com
Cmix: http://www.music.princeton.edu/winham/cmix.html
MAX/MSP: http://www.cycling74.com/products/dllmaxmsp.html

Freeware Real Rime Modular Synthesis tools
subsynth available at http://subsynth.sourceforge.net
ARTS available at http://www.arts-project.org

The MIDI protocol is perhaps the simplest method of implementing audio representations of data is using the MIDI protocol. This enables control of pitch, duration, loudness, set timbral variations, pitch bend and various other continuous controllers for up to 16 separate channels. MIDI messages are based on a simple 3 Byte system, components to convert these to MIDI synth signals can be downloaded at
http://www.bytamin-c.com/components/Component-MIDIsv121.htm
general info is available at http://midiworld.com/basics, detailed specification can be found at http://www.midi.com amongst many other places ...

Acknowledgments Thanks to Phil Husbands, Adrian Thompson and Chris Darwin for helpful comments on early manuscripts. Research is funded by EPSRC DPhil studentship.

References

[1] W.R. Ashby. Connectance of large dynamic (cybernetic) systems: Critical value for stability. *Nature*, 228(784), 1970.

[2] U. Axen and I. Choi. Investigating geometric data with sound. In S. Frysinger and G. Kramer, editors, *Proceedings of the Third International Conference on Auditory Display*, Palo Alto, CA, 1996. ICAD.

[3] S. Barrass and G. Kramer. Using sonification. *Multimedia Systems*, 7(1):23–31, 1999.

[4] F. Bastide. Representations of scientific practice. In S. Lynch, M. Woolgar, editor, *The iconography of scientific texts: principles of analysis*, pages 187–231, Cambridge, Mass, 1990. MIT Press.

[5] N. Birmbaumer, W. Lutzenberger, H. Rau, C. Braun, and G.M. Kress. Perception of music and dimensional complexity of brain activity. *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering*, 6(2):267–278, 1996.

[6] S.A. Bly. Communicating with sound. In *Proceedings of the CHI Conference on Human Factors in Computer Systems*, pages 115–119, New York, 1982. ACM.

[7] A.S. Bregman. *Auditory Scene Analysis*. MIT Press, Cambridge, MA, 1990.

[8] S.A. Brewster. Over coming the lack of screen space on mobile computers. *Tech report Department of Computing Science, University of Glasgow*, 2002.

[9] S.A. Brewster and M.G. Crease. Correcting menu usability problems with sound. *Behaviour and Information Technology*, 18(3):165–177, 1999.

[10] S.A. Brewster and R. Murray. Presenting dynamic information on mobile computers. *Personal Technologies*, 4(2):209–212, 2000.

[11] W. Buxton. Introduction to special issue on nonspeech audio. *Human Computer Interaction*, 4(1):1–9, 1989.

[12] F Dombois. Using audification in planetary seismology. In *Proceedings of the International Conference on Auditory Display*, pages 227–228. ICAD, 2001.

[13] D.G. Drew and T. Grimes. Audio-visual redundancy and TV news recall. *Communication Research*, 14:452–461, 1987.

[14] C. Earl and J.A. Leventhal. Survey of windows screen reader users: Recent improvements in accessibility. *Journal of Visual Impairment and Blindness*, 93(3), 1999.

[15] A.C. Eldridge. The effect of musical experience on preferred display modality for a cellular automata classification task. Technical report, School of Cognitive and Computing Sciences, University of Sussex, Brighton, UK, 2003.

[16] R. Findahl. The effect of visual illustrations upon perception and retention of news programmes, 1981. ERIC documentation reproduction services. no ED 054 631.

[17] T. Fitch and G. Kramer. Sonifying the body electric: Superiority of an auditory over a visual display in a complex, multi-variate system. In

Auditory Display: Sonification, Audification and Auditory Interfaces, Proceedings of the First International Conference on Auditory Display (ICAD), pages 307–326, Reading, MA, 1992. ICAD, Addison-Wesley.

[18] J. H. Flowers and D.C Grafel. Perception of sonified daily weather records. In *Proceedings of the Human Factors and Ergonomics Society*, pages 1579–1583, 2002.

[19] W.W. Gaver. The sonicfinder: an interface that uses auditory icons. *Human Computer Interaction*, 4(1):67–94, 1989.

[20] W.W. Gaver. What in the world do we hear? an ecological approach to auditory source perception. *Ecological Psychology*, 5(1):1–29, 1993.

[21] W.W. Gaver. Using and creating auditory icons. In G. Kramer, editor, *Auditory Display: Sonification, Audification and Auditory Interfaces*, volume XVIII of *SFI Studies in the Sciences of Complexity*, pages 417–446, Reading Mass, 1994. Addison Wesley.

[22] S. Handel. *Listening*. MIT Press, 1989.

[23] C. Hayward. Listening to the earth sing. In G. Kramer, editor, *Auditory display: Sonification, Audification and Auditory Interfaces*, volume XVIII of *SFI Studies in the Sciences of complexity*, Reading MA, 1994. ICAD, Addison Wesley.

[24] T Hermann, C Nieuhus, and H Ritter. Interactive visualisation and sonification for monitoring complex processes. In *Proceedings of the 2003 International Conference on Auditory display*, Boston, MA, USA, July 2003. ICAD.

[25] T Hermann and H Ritter. *Advances in Intelligent Computing and Multimedia Systems*, chapter Listen to your data: Model-Based Sonification for Data Analysis, pages 189–194. Baden-Baden, 1999.

[26] T Hermann and H Ritter. Crystallization sonification of high-dimensional datasets. In *Proceedings of the 2002 International Conference on Auditory display*, Kyoto, Japan, July 2002. ICAD.

[27] J.H Howard and J. A Ballas. Acquisition of acoustic pattern categories by exemplar observation. *Organisational Behaviour and Human Decision Processes*, 30(2):157–173, 1982.

[28] E. Jovanov and D. Starcevic. Acoustic rendering as support for sustained attention during biomedical procedures. In *Proceedings of the Fifth International Conference on Auditory Display*, Glasgow, UK, 1998. ICAD.

[29] A. Kennel. AudioGraph: A diagram reader for blind people. In *Proceedings of Second Annual ACM conference on Assistive Technologies*. ASSETS, 1996.

[30] N. Kitagawa and S. Ichihara. Hearing visual motion in depth. *Nature*, 416:172–174, 2002.

[31] G. Kramer, editor. *Auditory display: Sonfication, Audification and Auditory Interfaces*, volume XVIII of *SFI Studies in the Sciences of complexity*, Reading MA, 1994. ICAD, Addison Wesley.

[32] G. Kramer. Some organising principles for representing data with sound. In G.Kramer, editor, *Auditory display: Sonification, Audification and Auditory Interfaces*, volume XVIII of *SFI Studies in the Sciences of complexity*, pages 85–221, Reading MA, 1994. ICAD, Addison Wesley.

[33] G. Kramer, B. Walker, T. Bonebright, P. Cook, J. Flowers, N. Miner, and J. Neuhoff. Sonification report: Status of the field and research agenda. Technical report, National Science Foundation, 1997.

[34] D. Larsson, P. Vastfjall and K. Mendel. Ecological acoustics and the multi-modal perception of rooms: real experiences of auditory-visual virtual environments. In *Proceedings of the International Conference on Auditory Display*, pages 245–249. ICAD, 2001.

[35] P.A. Luca. An evaluation of the communicative ability of auditory icons and earcons. In *Auditory Display: Sonification, Audification and Auditory Interfaces, Proceedings of the Second International Conference on Auditory Display (ICAD)*, pages 307–326, Santa Fe, New Mexico, 1994. ICAD.

[36] D. Lunney and R. Morrison. High technologies for visually handicapped chemistry students. *Journal of Chemistry Education*, 58:228, 1990.

[37] G. Mawko and G. Kramer. Sonification for assessment of regional function in radionuclide ventriculography. proposal.

[38] K. McCabe and A. Rangwalla. Auditory display: Sonification, audification and auditory interfaces. In G. Kramer, editor, *Auditory Display of Computational Fluid Dynamics Data*, volume XVIII of *SFI Studies in the sciences of complexity*, pages 321–340, Reading Mass, 1994. Addison Wesley.

[39] H. McGurk and J. MacDonald. Hearing lips and seeing voices. *Nature*, 264(5588), 1076.

[40] R.D. Melara and L.E. Marks. Interaction among auditory dimensions: Timbre pitch and loudness. *Perception and Psychophysics*, 482:169–178, 1990.

[41] K.V. Nesbit and S. Barrass. Evaluation of a multimodal sonification and visualisation of depth of market stock data. In *Proceedings of the 2002 International Conference on Auditory display.* ICAD, 2002.

[42] J.G. Neuhoff. Perceptual bias for rising tones. *Nature*, 395:123–124, 1998.

[43] J.G. Neuhoff, G. Kramer, and J. Wayand. Sonification and the interaction of perceptual dimensions: Can the data get lost in the map? In P. Cook, editor, *Proceedings of the 2000 International Conference on Auditory Display*, Atlanta, Georgia, April 2000. ICAD.

[44] J.G. Neuhoff and J. Wayand. Pitch change, sonification and musical expertise: which way is up? In *Proceedings of the 2002 International Conference on Auditory display*, Kyoto, Japan, July 2002. ICAD.

[45] R.D. Patterson. Guidelines for auditory warning systems on civil aircraft. Technical report, Civil Aviation Authority, London, 1982.

[46] S.V. Pereverez, A. Loshak, S. Backhaus, J.C. Davis, and R.E. Packard. Quantum oscillations between two weakly coupled reservoirs of superfluid 3he. *Nature*, 388:449–451, 1997.

[47] H. Petsche, K. Lindener, P. Rappelsberger, and G. Gruber. The eeg: An adequate method to concretise brain processes elicited by music. *Music Perception*, 6:133–160, 1988.

[48] I. Pollack and L. Ficks. Information of elementary multidimensional auditory displays. *Journal of the Acoustical Society of America*, 26:155–158, 1954.

[49] K. Poppel. Temporal mechanisms in perception. *International Review of Neurobiology*, 37:185–202, 1994.

[50] D.A. Rabenhorst, E.J. Farrell, D.H. Jameson, T. D Linton, and J.A Mandelman. Complementary visualisation and sonification of multidimensional data. In *Proceedings of the SPIE Symposium on Electronic Imaging*, 1990.

[51] R. Sekuler, A.B. Sekuler, and R. Lau. Sound alters visual motion perception. *Nature*, 385:308, 1997.

[52] W.J. Severin. Another look at cue summation. *Communication Review*, 15:233–245, 1967.

[53] L. Shams, Y. Kamitani, and S. Shimojo. What you see is what you hear. *Nature*, 408:788, 2000.

[54] E. Tufte. *The visual display of quantitative information.* Graphics Press, 1992.

[55] J. Tzelgov, R. Srebro, and and A.Kushelevsky A. Henik. Radiation detection by ear and by eye. *Human Factors*, 29(1):87–98, 1987.

[56] P. Vickers and J.A. Alty. Musical program auralisation: Empirical studies. In P. Cook, editor, *Proceedings of the 2000 International Conference on Auditory Display*, Atlanta, Georgia, April 2000. ICAD.

[57] P. Vickers and J.A. Alty. When bugs sing. *Interacting with Computers*, 14(6):793–819, 2002.

[58] B.N. Walker and G. Kramer. Mappings and metaphor in auditory displays: an experimental assessment. In S. Frysinger and G. Kramer, editors, *Proceedings of the Third International Conference on Auditory Display*, pages 71–74, Palo Alto, CA, November 1996. ICAD.

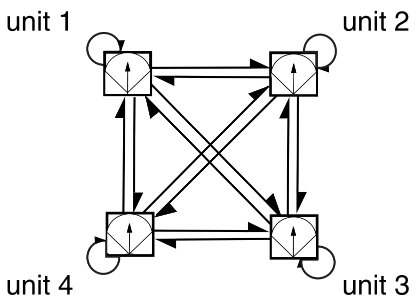
[59] B.N. Walker, G. Kramer, and D.M. Lane. Psychophysical scaling of sonification mappings. In P.Cook, editor, *Proceedings of the 2000 International Conference on Auditory Display*, Atlanta, Georgia, April 2000. ICAD.

[60] C.D. Wickens. *Varieties of Attention*, chapter Processing Resources in attention, pages 63–101. Academic Press, New York, 1984.

[61] C.D. Wickens, S.E. Gordon, and 1998 L.Liu. *An Introduction to Human Factors Engineering.* Addison Wesley Longmann, New York, 1998.

[62] S. Wolfram. Cellular automata as simple self-organising systems, 1982. Caltech preprint CALT - 68-938.

[63] D.N. Zotkin, R. Duraiswami, and L.S. Davis. Customizable auditory displays. In *Proceedings of the 2002 International Conference on Auditory Display*, Kyoto, Japan, July 2002. ICAD.



unit	output	pitch
1	0.00	C5
2	0.00	G#4
3	0.00	E4
4	0.00	C4

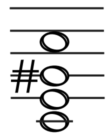


Figure 1: Diagram of connections for a fully connected 4 unit homeostat (left). Central value (0.00) of each unit is mapped to a discrete pitch value (right)

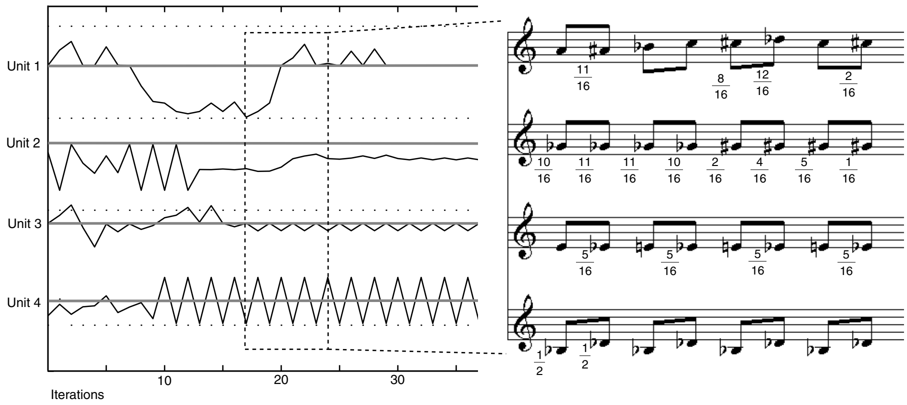


Figure 2: Graph of outputs of a 4 unit network (left), and corresponding auditory output of detail (right)

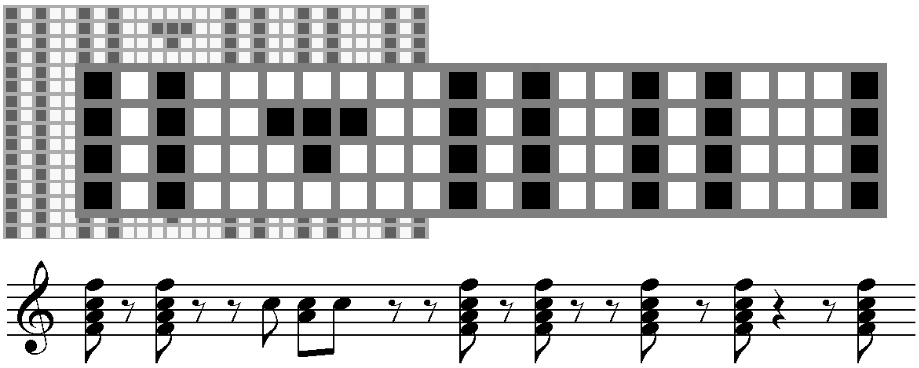


Figure 3: Example rhythmic and harmonic mapping for an ordered CA: binary state of each cell determines if state of note (On = play, Off = rest)

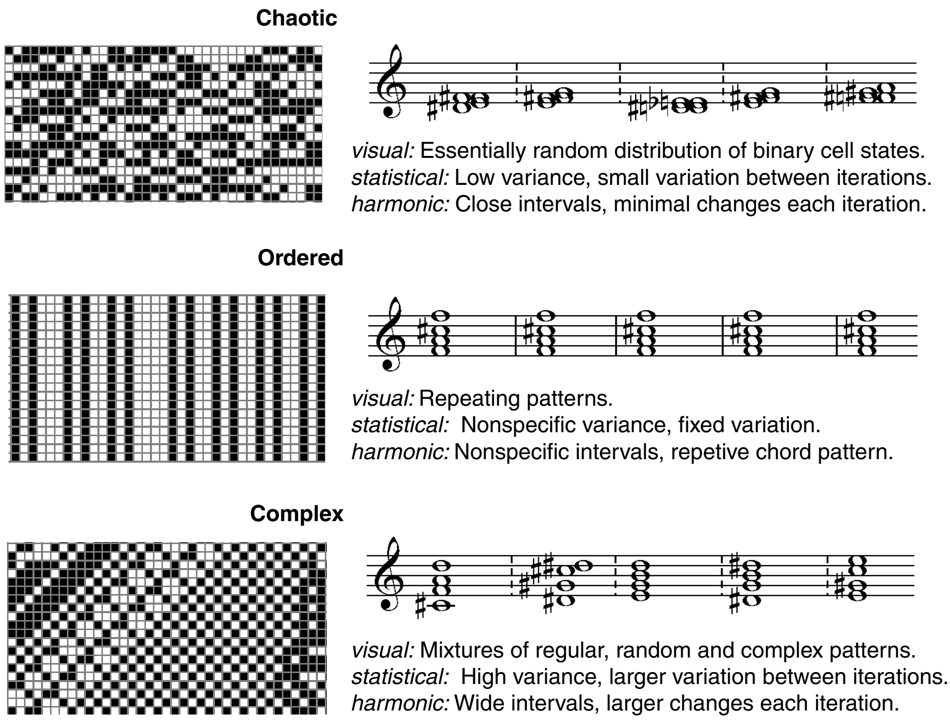


Figure 4: Examples of the harmonic representations of different CA types

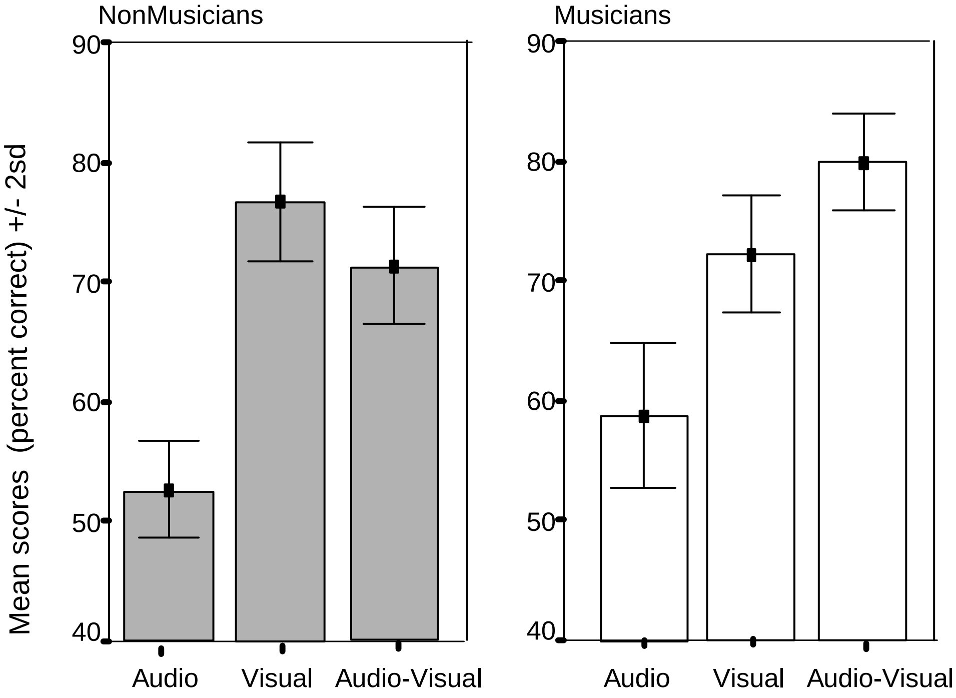


Figure 5: Mean scores and standard deviations in each condition, for both groups