PROTOTYPING A METHOD FOR THE ASSESSMENT OF REAL-TIME EEG SONIFICATIONS

Tony Steffert, Simon Holland, Paul Mulholland, Sheep Dalton.

Music Computing Lab,
Centre for Research in Computing,
The Open University, England
tony.steffert@open.ac.uk

Aleksander Väljamäe

Decision, Emotion and Perception lab
Dep. of Behavioural Sciences and Learning,
Linkoping University, Sweden
aleksander.valjamae@liu.se

ABSTRACT

This paper presents a first step in the development of a methodology to compare the ability of different sonifications to convey the fine temporal detail of the Electroencephalography (EEG) brainwave signal in real time. In EEG neurofeedback a person's EEG activity is monitored and presented back to them, to help them to learn how to modify their brain activity. Learning theory suggests that the more rapidly and accurately the feedback follows behaviour the more efficient the learning will be. Therefore a critical issue is how to assess the ability of a sonification to convey rapid and temporally complex EEG data for neurofeedback.

To allow for replication, this study used sonifications of pre-recorded EEG data and asked participants to try and track aspects of the signal in real time using a mouse.

This study showed that, although imperfect, this approach is a practical way to compare the suitability of EEG sonifications for tracking detailed EEG signals in real time and that the combination of quantitative and qualitative data helped characterise the relative efficacy of different sonifications.

1. INTRODUCTION

Electroencephalography (EEG) based neurofeedback (NF) training often requires subjects to monitor and track the frequency-specific band-power features of their own EEG, in order to learn how to modify some aspect of their brain activity. In some situations, such as relaxation training, the eyes may be closed or otherwise occupied on another task, in which case sound based feedback can be a valuable option.

Many neurofeedback practitioners consider that, due to established conditioning principles, [1, 2] the more faithfully and faster the band-power signal can be communicated in real time, the more effective the learning process is likely to be.

There has been a long history of using sonification to help understand the electrical activity of the human brain. In 1934, only 5 years after the neurologist Hans Berger first published his invention of the electroencephalograph, the Nobel laureate Prof. Edgar Adrian of Cambridge University reported the sonification of his own brain waves [3].

In the 1960s, Kamiya and colleagues [4, 5, 6] demonstrated some of the first examples of neurofeedback with the real-time auditory display of a participant's own alpha brain wave activity. In the mid-1960s, Sterman [7, 8] first with cats, and then with people with epilepsy, demonstrated the efficacy of neurofeedback. In the 1970s, Lubar [9] replicated Sterman's work with an ADHD population. In the 1990s Peniston and Kulkosky further developed neurofeedback relaxation procedures for alcoholism / posttraumatic stress disorder [10].

Due to technical constraints, many of the early studies used sound feedback exclusively, but in the 1990s with the development of computer technology and graphic displays in particular, most neurofeedback systems focused on visual displays, relegating sound feedback to a secondary role.

Sonification, with its objective, systematic and reproducible [11] sonic presentation of real-time EEG data, offers considerable untapped potential to convey many aspects of the rapid and complex nature of EEG in a manner accessible to the human auditory system.

Of the 100 or so EEG sonification papers in the last 80 years, around 70% have reported on sonifications that are capable of presenting the EEG data in real-time, and there have been a number of "proof of concept" papers that demonstrate interesting and novel approaches to EEG sonification [12]. Of these, 11 use sound for EEG neurofeedback [5, 6, 13-19]. However, to date there are only a few EEG sonification studies that provide a quantitative assessment of the ability of sonification to convey the character of EEG data and/or the listener's ability to perceive this physiological data. (For review see [20, 21].)

1.1. Motivation for this study

As the objective of neurofeedback is to learn how to modify one's own brain activity through feedback, the primary aim of this study is to develop a method to assess the efficacy of a sonification to convey EEG data in a manner that can facilitate the perception of the EEG data.

Although real-time feedback is critical in the neurofeedback loop, and the sonifications were specifically selected for this ability, in order to make a controlled comparison between the different sonifications, pre-recorded EEG was used.

In our intended eventual applications, subjects will only need to make a mental response to the sonifications, with no motor action necessary. However, in order to have some objective measure of how well the real-time changes in activity levels could be continuously perceived, participants were asked to track the activity in the sonification with a mouse. Clearly, having to make a motor response to the signal introduces a great deal of lag, and degrades performance. However this lag and degradation applies equally to all conditions and averaging or smoothing of the data to slow it down and make it more track-able is likely to reduce the information content and degrade perception of the finely detailed signal. In this study we combine a tracking task, workload questionnaire and subjective ratings of qualities such as the perceived arousal and valence of the sound.

Previously some EEG sonification studies have used the 'two-alternative forced-choice method' (2AFC) assessment method, where the participant is repeatedly presented with a sonification from one of two groups, e.g., patient with epilepsy versus a non-patient [22], or patients suffering from mild cognitive impairment versus healthy age-matched controls [23, 24, 25]. In such studies, after some initial training, participants are asked to pick which group a particular sonification file belongs to. Some of these studies have shown very good detection accuracy but this method does not really capture the temporal aspects in perception of the sonified data.

By contrast, some studies maintain some of the temporal information by getting participants to identify the onset of a particular EEG activity. So, for example, Khamis [26] played two channels of EEG sonification of patients with temporal lobe epilepsy and asked the study participants to push a button when they heard the onset of seizure activity. Khamis concluded "With only 2 h of training, non-expert subjects can detect seizures from audified EEG signals of 2 difference electrodes with a comparable degree of accuracy as can be done visually from review of EEG traces using the 10-20 electrode placements by an expert electroencephalographer".

From the point of view of assessing the temporal resolution of a sonification this is an improvement over the 2AFC method, but still does not capture the full range of dynamic characteristics of listening to continuous sound-based feedback. Furthermore, epileptic activity has a significantly larger amplitude and a very different morphology, which is very easily distinguished from the background EEG. Although this is an important area for applying EEG sonification, it is also somewhat specialised, since epilepsy only affects around one percent of the general population [27].

The development of a methodology (or set of methodologies) that could assess the ability of sonifications to convey in real-time, temporally rich EEG data would greatly assist the design and selection of appropriate sonifications for a range of application areas such as neurofeedback, surgical monitoring, or brain computer interfaces (BCIs) [28].

As an initial step in the development of a methodology to compare the effectiveness of real-time EEG sonification, the present study used an off-line listening method to compare two different sonification techniques. Both sonification algorithms are real-time-ready, but were applied to pre-recorded rather than live EEG data. This made it possible to control for variability in the signal, and allowed a within-subject study design, where all the participants heard the same sounds. The present study had two aims. Firstly to identify whether continuous real-time tracking of sonifications by means of a slider and computer mouse could be a practical assessment tool with non-expert users, and secondly, whether such a method could characterise the ability of sonifications to convey real-time EEG data.

2. MOTIVATION FOR CHOICE OF SONIFICATIONS

The electrical activity of the brain as measured from the scalp is a noisy, low amplitude signal (up to 100 microvolts), with a typical amplitude resolution of 1 microvolt, and a temporal resolution of around 2 milliseconds. EEG has complex temporal dynamics in the millisecond to decasecond time scale with a frequency range of 0 to over 70 Hz. The raw EEG has a distinct morphology and can be sub-divided into different frequency bands representing specific cognitive processes [29]. The EEG signal can be difficult to interpret, and neurologists and epileptologists usually specialise for several years to be able to interpret the EEG signal.

Although there are a large number of parameters that can be derived from the raw EEG including simple power, coherence, phase and complexity measures such as sample entropy, it is common in neurofeedback to measure and train only a few narrow band powers of the signal, such as Alpha activity. In a visual display, it is common to have a number of bar graphs that track up or down with the amplitude of the given band power. Because the EEG signal fluctuates so rapidly, the data driving the bar graphs is generally averaged over 1-second or even longer time windows, to reduce flicker and eye strain [20]. This averaging adds a delay, which may negatively affect the NF training progress [30].

In the design of sonifications to present EEG data, in order to maximize information transmission, perception and learning, a balance must be struck between converting as much of the complexity of the EEG data as possible into sound and between a person's ability to perceive and utilise the signal in the sound. By Hermann's definitions for sonification [11], the data transformation into sound must be objective, systematic and reproducible; at the same time, the purposes of neurofeedback require real-time sonification to render the time series data features in a salient, immediate, and contingent fashion [31].

To date, there has been a wide range of different data processing and sonification techniques used to display EEG, but few studies have tested sonifications against each other for their ability to convey the temporal dynamics of the EEG signal. This study is an initial step towards regularising methods for comparing EEG sonifications in the context of neurofeedback.

Audification is perhaps the simplest form of sonification mapping, in the sense that it simply maps the input data to sound pressure levels. This could be thought of as the auditory equivalent of looking at a raw EEG trace. But, bearing in mind that 98% of the EEG power is below 30 Hz [29], naive real-time audification would produce results below the human auditory range. Thus, a carrier wave is needed, bringing us to a contrasting approach to sonification: amplitude modulation.

Amplitude Modulation (AM) sonification could be seen as analogous to the bar graph of a band power used in a typical neurofeedback display, as the power of EEG band increases, the bar graph goes up and so does the volume of the sound. Conceptually AM sonification is simple (though this is no guarantee of perceptual simplicity). But despite the simplicity, it is not obvious how well this mapping might allow listeners to track rapid level changes of the kind typical of EEG - this is a matter to be established empirically.

Frequency Modulation (FM) sonification maps changes in the input signal to changes in the frequency of the sound output. Frequency has obvious potential for communicating relatively rapid and fine changes in real time, but again, it is unclear how well this mapping might be suited to our particular purposes.

Because of the mapping simplicity of AM and FM, the only subjective design decisions needed are to select the carrier wave frequency for the AM sonification, and the output frequency range for the FM sonification. Both can be readily chosen to fit comfortably within the human auditory frequency range.

Thus, Amplitude Modulation (AM) and Frequency Modulation (FM) sonifications were the first two continuous data representations [32] parameter mapping [11] methods chosen for comparison. By starting with these conceptually simple, easy-to-generate sonifications that require a minimum of subjective design decisions, the intention is to establish an initial baseline measure. Such a measure has the potential to facilitate comparison of more complex and engaging sonifications. Use of open source research presentation and sound synthesis software allows straightforward reuse by other researchers.

At the outset we were confident in predicting that Frequency Modulation sonification would outperform Amplitude Modulation at conveying the real-time EEG data and allowing for more accurate tracking of the sonified alpha band envelope. However, this left plenty of room for surprises: since the study uses both subjective and objective tracking measures, ample scope remained for contrasts between findings: for example, a better tracking score could come from a sonification with a worse subjective rating in terms of task load, emotional ratings of valence and arousal, or aesthetic quality of the sound.

3. METHODS

3.1. Electroencephalogram measurement and processing

Six, 3 minute, 19 channels "Full Cap" EEGs were recorded in two conditions, eyes closed and eyes open, using the first author as a subject. The EEG was recorded with a Mitsar 202 amplifier and WinEEG software [33] at a sample rate of 2000 Hz and saved at 500 Hz, in a linked ears referential montage. The low cut filter was set to 0.53 Hz and the high at 50 Hz, the notch filter was 45 to 55 Hz and all impedances were kept below 5 kilohms. In Matlab 11b [34] the EEG was band-pass filtered with Butterworth IIR filter of order 5, to make two EEG bands, one of low alpha (LA) 7-10 Hz, and the other of high alpha (HA) 10-13 Hz and the Hilbert transform was used to extract the amplitude envelopes of alpha EEG signals.

Alpha activity generally increases when sensory information is reduced to the brain. For example, when the eyes are closed, more alpha is produced in the occipital cortex in the back of the head. Consequently, the 'eyes closed' condition is typically a lower arousal state than 'eyes open' and generally has more alpha activity in most people [29]. Traditionally, alpha has been defined as a band of 8 to 12 Hz, but newer research suggests that the upper and lower alpha bands represent different cognitive functions [35]. The electrode location Pz in the back of the head was selected because it has a good level of alpha activity and is commonly used in neurofeedback for relaxation training.

By visual examination of the raw alpha signal and spectral content of all of the EEG files, four 1 minute files were selected that captured a selection of typical alpha activity in eyes closed and eyes open and in the High and Low Alpha frequency conditions. In the remained of this section, we consider the characteristics of these four sample EEG files used for the study, as summarised in table 1.

In table 1, the names of the EEG files are; 'HAO' is the high alpha band in the eyes open conditions state. 'HAC' is high alpha with eyes closed. 'LAO' is low alpha with eyes open and 'LAC' is low alpha with eyes closed.

The contents and meaning of the various columns in Table 1 are as follows: 1) the number of alpha bursts, quantified as alpha activity over the grand mean for longer than 280 ms; 2) the mean duration of the alpha bursts in seconds; 3) excess kurtosis of the alpha amplitude envelopes (which is a measure of the pointedness or flatness of the histogram of the distribution - the smaller the number, the closer to a normal distribution and the less pointed the peak - negative values indicate flatness of the peak); and 4) the skewness (which is a measure of how symmetrical the distribution of the data is around the mean, and the distribution of the 'tails').

Considering table 1 overall, although there is a clear visual difference in these sample files in the patterns of alpha amplitude envelope activity between the eyes open and eyes closed conditions, the number of alpha bursts and the mean duration do not show a large difference.

The eyes-open alpha EEG had a high excess kurtosis distribution (i.e. high peakedness or leptokurtic) and is more positively skewed, compared to the eyes closed EEG, suggesting the eyes open EEG has fewer and shorter large amplitude "bursts". The eyes closed alpha EEG was closer to a normal distribution on both kurtosis and skewness with a flatter peak of distribution implying more mid-range activity.

Table 1: *Quantification of alpha activity in four EEG files*

	# of alpha bursts	Mean duration alpha bursts [s]	Excess kurtosis	Skew
HAO	35	0.60 [0.33]	0.75	0.98
HAC	40	0.57 [0.38]	-0.22	0.37
LAO	38	0.51 [0.19]	3.15	1.42
LAC	35	0.58 [0.31]	0.23	0.69

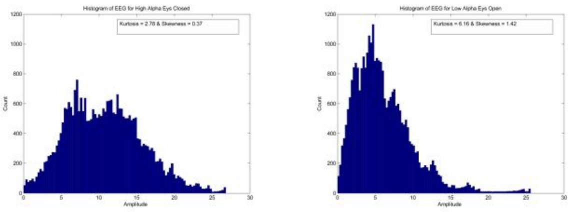

Figure 1: *Histograms of the EEG alpha activity: the High Alpha, Eyes Closed (left panel) and the Low Alpha Eyes Open (right panel).*

3.2. Sonification of EEG

Alpha signal envelopes were imported into Pure Data software [36] where any EEG values greater than 30 microvolts (mV) were set to 30 mV, to exclude artifacts like eye blinks and muscle tension, so that the data values ranged between 0 and 30 mV. The audio sample rate was set to 48,000 Hz and the four 1 minute EEG files were sonified with AM and FM-based methods. Two different audio frequency outputs were chosen for the carrier wave to control for any bias in the hearing or aesthetic response of the participants. Each carrier frequency was presented in 4 sound files to counterbalance across conditions of eyes open and closed and high and low alpha.

For AM sonification, each data point was divided by 30 to scale the values to range between 0 and 1. The data was then linearly interpolated to match the EEG to the audio sample rate. Half of the files were then multiplied by a sine wave carrier of either, 261.6 Hz (Middle C) or 523.2 Hz and the output saved as a .wav file. These two carrier frequencies from an idealized critical band filter bank [37] were used (3rd and 6th band) in order to compare the results from this current experiment with a future planned study where two streams will be presented simultaneously.

For FM sonification the EEG data was multiplied by a factor of 20 to give an output range of 0 to 600 and then each value was added to by either 261.6 or 523.2, giving an output frequency range of 261.6 to 861.6 Hz or from 523.2 to 1123.2 Hz. The output was then linearly interpolated to audio sample rate and saved as a .wav file.

3.3. Participants, experimental design and procedure

Seventeen participants, mean age 45.65 (SD = 13.09), 8 females, took part in the experiment. All had a normal level of vision, hearing and cognitive functioning and were over 18 years old. The participants signed a consent form, were not paid or given any inducements to participate and were informed they had the right to withdraw at any time and their data would be destroyed. The study received ethics approval from the Open University Human Research Ethics Committee number HREC/2014/1733/Steffert and was conducted in accordance with the Declaration of Helsinki [38].

Participants were seated in front of a laptop with Sennheiser HD 439 Headphones on and played some example sounds to set the volume and practice the tracking task. All stimuli and questionnaires were presented using PsychoPy [39], an open source presentation software tool.

Participants were asked to track the activity of the sonification with a horizontal slider on the computer screen using the mouse. For the AM sonification, participants were instructed that they should move the slider to the right as the volume of the sound increased and to the left as it decreased. For the FM sonification the instruction was the same but for frequency.

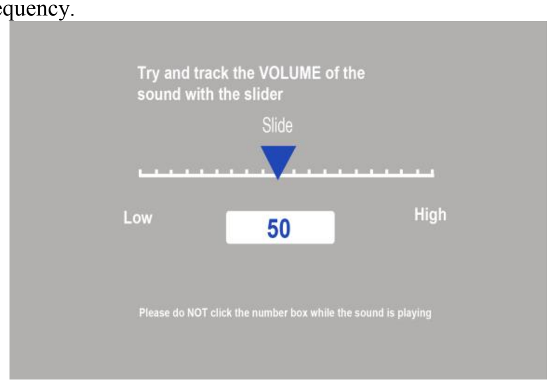

Figure 2: *Example of the Tracking Screen in PsychoPy.*

The goal of the tracking task is to test the whole data chain, from the data's transformation into sound, to the sound's conversion into perception and perception into a motor response of the participant. The testing session took between 15 and 25 minutes. 8 stimuli were used comprising of 2 (FM vs. AM) x 2 (eyes closed/eyes open) x 2 (Low Alpha vs. High Alpha) design. The presentation order was randomized across participants.

After listening to each sound file the participants were asked to rate on a 20 point Likert type scale both the arousal and valence [40] of the sound (the screen was similar to the tracking screen seen in figure 2). The arousal question was "How exciting/energetic or passive/relaxing was the sound?" and the Valence question was "How positive/happy or negative/sad was the sound?" The left side of the slider was marked either "passive/relaxing" or "negative/sad" and scored 1 while the right side was marked "exciting/energetic" or "positive/happy" and scored 20.

Then participants were asked how easy or difficult they found the tracking task with the six questions from the NASA-TLX workload questionnaire: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration [41]. The questions were presented in a random order on each trial and the participant had to rate the questions with a slider with "low" on the left that scored 1 to "high" on the right with a score of 20, except for the 'performance' rating that ranged from "good" on the left to "poor" on the right.

Questions about age, gender and musical experience were left till the end of the study to minimize stereotype threat [42] which is the participant perception of the researcher's expectation, which has been shown to affect performance. The four questions to assess the musical experience were: M1) "I engaged in regular, daily practice of a musical instrument (including voice i.e. singing) for "X" or more years", M2) "At the peak of my interest, I practiced "X" or more hours per day on my primary instrument", M3) "I have had "X" or more years of formal training on a musical instrument (including voice) during my lifetime", and M4) "I have had formal training in music theory for "X" or more years".

In a short post-experimental interview the participants were asked two questions: "Did these sounds remind you of any sound?" and "What do you think brainwaves would sound like if you could hear them?"

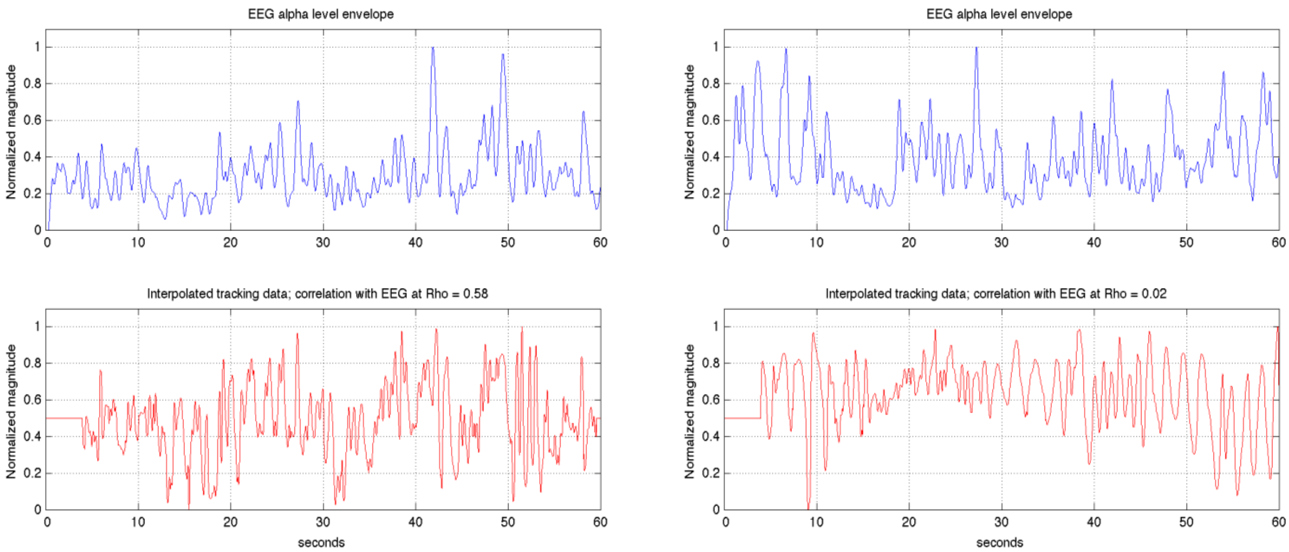

Figure 3: *EEG alpha level envelopes that were used for sonification and corresponding interpolated tracking data. Left panel – good tracking example (Rho = 0.58), Right panel – bad tracking example (Rho = 0.02). First 4 s of tracking data are replaced by constant value since this data was changed by the spline function.*

3.4. Tracking data pre-processing

Tracking accuracy was computed by correlating the original EEG data that generated the sonification with the participants' slider response to the sound output of the sonification.
First, the tracking data points (from each time the slider changed the position) was interpolated using cubic spline data interpolation in Matlab, to match the time scale and sampling rate (500Hz) of the EEG data. EEG data was also pre-processed by extracting the amplitude envelope using a Hilbert transform, and then using moving average window of 200 sample length (0.4 s). To compensate for differences in participant's reaction time and therefore variations in the lag of the tracking data, an iterative process to compute the correlation coefficient for all delays of up to 1 second to find the maximum was implemented in Matlab. The best match was also visually inspected to minimise the risk of erroneous matches (see Figure 3).

4. RESULTS

The Alpha level was fixed at 0.05 for all statistical tests. Greenhouse-Geisser correction was used to correct for unequal variances. For multivariate analysis Wilks' Lambda L was used as the multivariate criterion. All variables were normally distributed according to the Kolmogorov-Smirnov test. As there were no significant differences between low and high frequency alpha sonifications for any measure, they were combined for subsequent analysis.

The mean "tracking accuracy" i.e. the Pearson correlation coefficient Rho between the EEG data and the tracking data ranged between 0 and 0.58 (SD = 0.2). For seven participants the max correlation coefficient for all 8 conditions was lower than 0.4. As this is somewhat low, this suggests that some of the participants could either, not hear the signal in the sonification, or could not move the slider very accurately to track the data, or both.

A two-way within-subjects MANOVA was conducted using the 6 questions from NASA-TLX, subjective emotional ratings of valence and arousal (VAL and ARO), and 'Tracking accuracy' correlation coefficient Rho. The design was sonification type (FM/AM) x EEG condition (eyes closed/eyes open).

Four questions regarding musical experience were used for creating 2 types of subgroups. The first type was based on answers from M1 and M2 questions and forming subgroups with (10 out of 17 participants) and without musical instrument experience. The second type was based on answers from M3 and M4 questions and forming subgroups with (10 out of 17 participants) and without formal musical education. The two resulting groupings regarding musical experience differed slightly from each other (by 4 people).

4.1. The difference between AM and FM sonifications

The overall multivariate effect of sonification type was significant, with the difference between AM and FM at Wilks' Lambda = .108, $F(9, 8) = 7.34$, $p < .005$, $\eta^2 = 0.892$. Univariate tests showed significance of this modulation type effect for a number of measures. For the **Mental Demand** scale, difference was at $F(1, 16) = 7.05$, $p < .05$, $\eta^2 = 0.306$, showing that FM was reported as having higher mental demand than AM-based sonification, (M =11.2 SD = 1.2) vs. (M = 9.4 SD = 1.1). For the **Physical Demand** scale the significance was at $F(1, 16) = 8.66$, $p < .01$, $\eta^2 = 0.351$, with FM being reported as requiring more physical activity (M = 7.6, SD = 1.2) than AM-based sonification (M = 5.8, SD = 0.8). For the **Temporal Demand** scale the significance was at $F(1, 16) = 7.45$, $p < .05$, $\eta^2 = 0.318$, with FM-based sonification being rated as having more time pressure (M = 10.9, SD = 1.4) than for AM-based (M = 8.3, SD = 1.0). For the **Effort** scale the difference was significant at $F(1, 16) = 9.3$, $p < .01$, $\eta^2 = 0.368$ with FM requiring greater effort (M = 10.7, SD = 1.3) than AM-based sonification (M = 8.7, SD = 1.2). On the subjective **arousal** scale, FM-based sonification was significantly more exiting/energetic (M = 12.8, SD = 1.1) than AM-based one (M = 8.1, SD = 0.8) with $F(1, 16) = 24.49$, $p < .001$, $\eta2 = 0.605$. Finally, for the **tracking accuracy** the Rho values were significantly higher for FM-based (M = 0.21, SD = 0.34) than for AM-based sonification (M = 0.13, SD = 0.36) at $F(1, 16) = 9.92$, $p < .01$, $\eta^2 = 0.383$.

On a few scales, differences between two sonifications could be observed, but they did not reached significance. For the **valence** scale, the difference between FM and AM sonification was close to significance with $F(1, 16) = 3.18$, $p = .1$, $\eta^2 = 0.166$ with FM being judged more positive/happy (M = 9.4 SD = 1.0) than AM (M = 7.9 SD = .7). **Frustration** was higher for FM (M = 10.4, SD = 1.1) than for AM (M = 9.36, SD = 1.0) but did not reach significance $F(1, 16) = 2.42$, $p = .14$, $\eta^2 = 0.131$. Interestingly, despite FM being rated higher than AM on all the other measures, the self-rating of **Performance** showed no difference between the two sonification methods. The difference was at $F(1, 16) = .302$, $p = 0.59$, $\eta^2 = 0.019$, with FM (M = 10.41, SD = 1.2) and AM (M = 9.96, SD = 1.1) on a scale of 1 to 20.

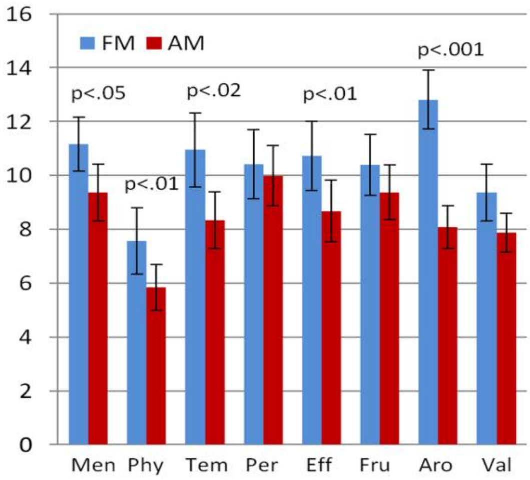

Figure 4: *The vertical axis shows the mean and Standard Error of the subjective ratings on a 20 point Likert type scale ranging from 1 to 20 for the six question of the NASA-TLX: Mental Demand (Men), Physical Demand (Phy), Temporal Demand (Tem), Performance (Per), Effort (Eff), Frustration (Fru), as well for Arousal (Aro) and Valence (Val), with the p-values for the statistically significant differences between FM in (blue) and AM (red).*

Although the tracking accuracy was significantly lower than in earlier pilot testing, and nearly all participants reported difficulties in moving the slider fast enough to keep up with the sound, the combination of continuous tracking data and subjective work load assessments of the tracking task has provided some interesting insights, as will now be summarised.

Overall the 17 participants performed better on tracking the FM sonification than the AM, but did not feel their performance was any better. They found tracking of FM sonification more mentally, physically and temporally demanding and more effortful but did not feel any difference in frustration between the two sonifications.

This could be interpreted as indicating that the participants could hear the data more accurately with the FM sonification therefore performed the tracking task more accurately and as a consequence of hearing more information, found the task more demanding. In other words, those who did not perceive the modulation may have found the task "easy" because they were unaware they were missing data and therefore found the task less demanding. This interpretation seems to agree with some previous non-EEG sonification studies [43] suggesting that FM sonification is generally better than AM sonification for presenting data.

4.2. The effect of the EEG condition

Participants rated the sonifications of EEG from eyes closed condition as having a higher Frustration ($M = 10.75$, $SD = 1.1$) than the eyes open condition ($M = 9.02$, $SD = 1.0$) with a statistical significance $F(1, 16) = 6.15$, $p = 0.025$, $\eta^2 = 0.278$, regardless of sonification type or frequency band. This may be because there is more alpha activity in the eyes closed condition with more variability. No interaction between EEG and sonification type reached significance.

4.3. Musical experience

Ten out of 17 participants had musical experience either in the form of playing an instrument or some formal training, music theory training and practiced at least 30 minutes a day at some time in their life. Hence, we created two grouping factors and repeated the two-way within-subject MANOVA with additional grouping factor of either musical instrument experience, or musical education.

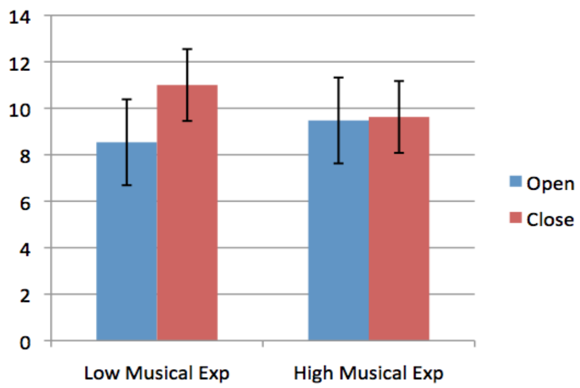

Figure 5: *Interaction between temporal demand factor from NASA-TLX and subgroups of musical experience level (playing any musical instrument or not). The open and close legend stands for sonification of EEG data from open or closed eyes condition.*

No significant effect for the musical education factor was found. However, two significant interactions between musical instrument experience and stimuli type could be seen. First, sonification type interacted with subgroups factor for the arousal ratings at $F(1, 16) = 5.33$, $p = 0.036$, $\eta^2 = 0.262$. Those who played a musical instrument found FM sonification a lot more arousing ($M = 14.13$, $SD = 1.4$) than those that did not ($M = 10.93$, $SD = 1.7$), but there was not such a big difference between subgroups in the arousal ratings to the AM sonification (8.5 vs. 7.8). Second significant interaction could be seen between the subgroups and EEG condition at $F(1, 16) = 5.59$, $p = 0.032$, $\eta^2 = 0.272$. The participants that did not play a musical instrument found that the sonification of EEG from eyes closed condition ($M = 11.00$, $SD = 1.9$) were more temporally demanding to track, that from the eyes open condition ($M = 8.54$, $SD = 1.8$). No such difference could be seen for listeners with music experience.

4.4. Post-experimental interviews

To the question "Did these sounds remind you of any sound?" Only one person said "No" and the largest answer with 6 (27 %) said it reminded them of "wind". Two people thought the sonifications sounded like "The Clangers" from the UK Children's TV show and most of the other answers shared a similar theme - replies included; "police siren" "vacuum cleaner machine", "whistle", "trombone", oscilloscope", "AV meter" and "happy complaining ghosts". Some people did not like the sounds at all and said it reminded them of "horror movies" or sounded like a "cheese grater".

For the question "What do you think brain waves would sound like if you could hear them?" Two people did not answer, three said "wind" (16%) and two thought the sonification did sound like brainwaves (11%) and 8 (42%) of the responses had a theme of busy activity like "boiling water", "busy like a switch board", "a terrible rowing noise" and "noise, white noise". One person said "music" and another "like a cheese grater".

5. DISCUSSION

The present study can be seen as an initial step in the development of a methodology to compare the effectiveness of real-time EEG sonifications. The main finding of the listening tests of 17 participants was that despite the tracking of FM sonification being rated as more mentally, physically and temporally demanding and taking more effort, the continuous tracking accuracy was significantly more accurate than for AM. Nearly 90% of the variability in combined measures comparison (MANOVA) can be explained by the type of sonification (i.e. FM or AM). Importantly, without a quantitative behavioural measure of a person's ability to perceive the data changes, the results of subjective evaluation would lead to the false conclusion that the AM sonification was a better method as it was rated as easier to track.

Only a few participants liked the wailing sounds of the AM and FM sonifications and some vehemently disliked them and three participants came close to terminating their involvement in the study. Despite the conceptual simplicity of the sounds, many participants either thought the sonifications sounded like brain waves or had some similarities to what they expected brain waves to sound like.

This study used pre-recorded EEG fragments that captured a range of different alpha activity patterns that exemplified the typical activity of eyes closed and eyes open conditions. However, there was only one statistically significant difference between sonifications of EEG from eyes closed and eyes open condition: participants rated data from eyes closed condition more frustrating to track. Interestingly, when adding musical experience as a subgroup variable, we saw that listeners who do not play any musical instrument found EEG sonification of the eyes closed data significantly more temporally demanding to track as compared to their own ratings of eyes open sonification, and to the ratings from users with musical experience. But it should be remembered that the 6 questions from the NASA-TLX were about the *workload* of the tracking task and only the arousal and valence ratings were about the *quality of the sound* of the sonifications.

This highlights a distinctive feature of this study, which used continuous real-time tracking to measure the difference in trackability between two types of sonification, without using sonification to identify or sort the data. The study also contrasts with those that solely measure subjective preferences for sonifications. As previously noted, there are a few EEG sonification studies that use the 'two-alternative forced-choice method' and some identify the onset of a particular activity. But one of the shortcomings of such methodologies is their inability to assess the temporal dynamics of the data and its perception.

The field of psychoacoustics has been researching sound and music perception for over one hundred and fifty years, so methodologies from this domain may help to illuminate the present study. But one of the problems with many psychoacoustic studies is that they tend to use very short sound clips that may not capture the temporal dynamics of a typical sonification listening session. So, for example, the International Affective Digitized Sounds (IADS) [44], which has created a normative emotional stimuli database, uses sounds of only 6-seconds in duration.

On the other hand, administering a questionnaire at the end of a 1 to 5-minute listening epoch will also fail to capture the temporal dynamic nature of most sound/music. Madsen [45] argues that what is needed is a "continuous non-verbal measurement of a participant's response to the music/sound stimuli that can expose the dynamic contours of a listening experience without distracting the participant from the listening task". To this end, Madsen and colleagues at the Center for Music Research at Florida State University have developed and validated with a large number of studies a 'Continuous Response Digital Interface' [46] that allows the user to turn a dial in real-time to log their immediate and continuous response on a continuum between two extremes such as "Positive" to "Negative" or "Lively" to "Passive". This current study could be seen as a variant of the Madsen methodology but within the sonification domain.

The objective of this research was to develop a sonification validation method that is specifically suited to the nature of real-time EEG feedback as opposed to time series data in general.

Whilst the continuous tracking of a sound stream with a mouse is a poor proxy for the perceptual decoding of a continuous signal, any lag from the motor response will apply equally to all conditions, and this study has shown that such an approach can generate a quantitative assessment of the real-time trackability of a sonification. Furthermore, although some older users without computer experience had difficulties tracking, and despite considerable variability in tracking accuracy between participants, the combination of quantitative and qualitative data helped to illuminate the relative usefulness of each sonification method.

It is hoped that by using the Open Source PsychoPy software platform for this tracking task and by establishing a web repository where sample EEGs, sonification code, sound files and presentation software can be collected and made freely available, there is the potential for future EEG sonification studies to add to a database of quantitative assessment of sonifications. Such a store could become a valuable resource for the development of the field.

Subsequent studies will include the simultaneous sonification of more data features, such as multiple channels of EEG, multiple frequency components, and multiple statistics of the same data feature. Crucially future work will compare real-time sonifications of EEG in the context of neurofeedback, as well as sonifying two peoples EEG at the same time so they can learn from each other [47].

6. ACKNOWLEDGEMENTS

The last author of this paper received funding from Marie Curie Actions of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA GA-303172. The authors would like to thank Andrey Bilichenko for help in prescreening and selecting EEG data [47].

7. APPENDIX

The EEGs and sound files as well as the PD and PsychoPy scripts can be found at http://www.sonification.qeeg.co.uk/

8. REFERENCES

[1] B. F. Skinner, Science and Human Behaviour. Pearson Education, Inc, p. 480, 1953.

[2] B. M. Sterman, "Physiological origins and functional correlates of EEG rhythmic activities: Implications for self-regulation.," *Biofeedback Self. Regul.*, vol. 21, no. 1, pp. 3–33., 1996.

[3] E. D. Adrian and B. H. C. Matthews, "The Berger rhythm: potential changes from the occipital lobes in man," *Brain*, vol. 57;, no. 1, pp. 355–385., Jan. 1934.

[4] J. Kamiya, "Conditional discrimination of the EEG alpha rhythm in humans," in *Western Psychological Association*, 1962.

[5] D. P. Nowlis, and J. Kamiya, "The control of electroencephalographic alpha rhythms through auditory feedback and the associated mental activity", *Psychophysiology*, 6(4), pp. 476–484, 1970

[6] J. V. Hardt and J. Kamiya, "Anxiety Change Thought Electroencephalographic Alpha Feedback Seen Only in High Anxiety Subjects". *Science,* vol. 201 no. 43, pp. 79–81, 1978.

[7] B. M. Sterman, "Suppression of seizures in an epileptic following sensorimotor EEG feedback training", *Electroencephalogr Clin Neurophysiol*, vol. 33, no. 1, pp. 89–95., 1972.

[8] B. M. Sterman, "Basic Concepts and Clinical Findings in the Treatment of Seizure Disorders with EEG Operant Conditioning", *Clin. Electroencephalogr.*, vol. 31, no. 1, 2000.

[9] J. F. Lubar and M. N. Shouse, "EEG and behavioral changes in a hyperkinetic child concurrent with training of the sensorimotor rhythm (SMR)" *Biofeedback Self. Regul.*, vol. 1, no. 3, pp. 293–306, Sep. 1976.

[10] E. G. Peniston and P. J. Kulkosky, "alpha-theta Brainwave Training and beta-Endorphin Levels in Alcoholics" *Alcohol. Clin. Exp. Res.*, vol. 13, no. 2, pp. 271–279, Apr. 1989.

[11] T.Hermann, "Taxonomy and Definitions for Sonification and Auditory Display". In: Susini P, Warusfel O, eds. *Proceedings of the 14th International Conference on Auditory Display (ICAD 2008)*. Paris, France: IRCAM; 2008.

[12] T. Hermann, G. Baier, U. Stephani, and H. Ritter, "Vocal sonification of pathologic EEG features.," in

Proceedings of the International Conferenceon Auditory Display (ICAD 2006), 2006.

[13] J. J.Allen, E. Harmon-Jones, & J. H.Cavender, “Manipulation of frontal EEG asymmetry through biofeedback alters self-reported emotional responses and facial EMG”, *Psychophysiology*, 38, pp. 685-693. 2001.

[14] J.Fell, H. Elfadil, P.Klaver, J. Röschke , C. E. Elger, and G.Fernández, “Covariation of spectral and nonlinear EEG measures with alpha biofeedback”, The International journal of neuroscience, 112(9), pp. 1047–57, 2002.

[15] T.Hinterberger, “The Sensorium: A Multimodal Neurofeedback Environment”, Advances in Human-Computer Interaction, 2011, pp. 1–10, 2011.

[16] S. W. Choi, S. E. Chi, S. Y. Chung, J. W. Kim, C. Y. Ahn, and H. T. Kim, “Is alpha wave neurofeedback effective with randomized clinical trials in depression? A pilot study”, *Neuropsychobiology*, 63(1), pp. 43–51, 2011.

[17] G. J. M. Van Boxtel, A. J. M. Denissen, M. Jäger, D. Vernon, M. K. J. Dekker, V. Mihajlović, et al., “A novel self-guided approach to alpha activity training”, *International journal of psychophysiology*, 83(3), pp. 282-94, 2012.

[18] J. V. Hardt, “Alpha brain-wave neurofeedback training reduces psychopathology in a cohort of male and female Canadian aboriginals”, *Advances in mind-body medicine*, 26(2), pp. 8–12, 2012.

[19] S. Wang, Y. Zhao, S. Chen, G. Lin, P. Sun, and T. Wang, “EEG biofeedback improves attentional bias in high trait anxiety individuals”, *BMC neuroscience*, 14(1), p. 115, 2013.

[20] A. Väljamäe, S. Mealla, T. Steffert, S. Holland, X. Marimon, R. Benitez, et al., “A review of real-time EEG sonification research”, in *International Conference on Auditory Display*, pp. 85-93, 2013.

[21] G. Dubus and R. Bresin, “A systematic review of mapping strategies for the sonification of physical quantities”, *PLoS One*, vol. 8, no. 12, p. e82491, 2013.

[22] P. Loui, M. Koplin-Green, M. Frick, and M. Massone, “Rapidly learned identification of epileptic seizures from sonified EEG”, *Front. Hum. Neurosci.*, vol. 8, no. October, p. 820, Jan. 2014.

[23] F. B. Vialatte, T. Musha, and A. Cichocki, “Sparse Bump Sonification: a New Tool for Multichannel EEG Diagnosis of Brain Disorders”, Artificial Intelligence in Medicine, 2009.

[24] D. Wu, C.-Y. Li, and D.-Z. Yao, “Scale-free music of the brain”, *PLoS One*, vol. 4, no. 6, p. e5915, 2009.

[25] F. B. Vialatte, J. Dauwels, T. Musha, and A. Cichocki, “Audio representations of multi-channel EEG: a new tool for diagnosis of brain disorders”, *Am. J. Neurodegener. Dis.*, vol. 1, no. 3, pp. 292–304, 2012.

[26] H. Khamis, A. Mohamed, S. Simpson, and A. McEwan, “Detection of temporal lobe seizures and identification of lateralisation from audified EEG”, *Clin. Neurophysiol.*, vol. 123, no. 9, pp. 1714-20, 2012.

[27] D. J. Thurman, E. Beghi, C. E. Begley, A. T. Berg, J. R. Buchhalter, D. Ding, D. C. et al, “Standards for epidemiologic studies and surveillance of epilepsy”, *Epilepsia*, vol. 52 Suppl 7, pp. 2–26, Sep. 2011.

[28] E. Curran, “Learning to control brain activity: A review of the production and control of EEG components for driving brain–computer interface (BCI) systems”, *Brain Cogn.*, vol. 51, no. 3, pp. 326–336, Apr. 2003.

[29] J. D. Kropotov, *Quantitative EEG, Event-Related Potentials and Neurotherapy*. Academic Press Inc, 2010, p. 600.

[30] T. H. Budzynski, H. K. Budzynski, J. R. Evans, and A. Abarbanel, *Introduction to Quantitative EEG and Neurofeedback: Advanced Theory and Applications*. Academic Press, 2009, p. 528.

[31] Collura, T.F., *Technical Foundations of Neurofeedback*. New York: Routledge / Taylor & Francis, 2014

[32] A. De Campo, “A Data Sonification Design Space Map”, in *Proceedings of the 2nd International Workshop on Interactive Sonification, York, UK* February 3, pp. 1–4, 2007.

[33] Mitsar 202 & WinEEG Software, Mitsar Medical, St. Petersburg, Russia.

[34] Matlab, The MathWorks, Inc. Natick, Massachusetts, United States. 2011.

[35] W. Klimesch, M. Doppelmayr, H. Russegger, T. Pachinger, and J. Schwaiger, “Induced alpha band power changes in the human EEG and attention”, *Neurosci. Lett.*, vol. 244, no. 2, pp. 73–6, Mar. 1998.

[36] M. Puckette, “Max at Seventeen,” *Comput. Music J.*, vol. 24, no. 4, pp. 31–44, 2002.

[37] B. Scharf, “Critical Bands”, in *Foundations of Modern Auditory Theory*, J. V. Tobias, Ed. San Diego: Academic, 1970, pp. 159–202.

[38] World Medical Association, “Declaration of Helsinki: Adopted by the 18th World Medical Assembly, Helsinki 1964 as amended by the 52nd World Medical Assembly, Edinburgh, Scotland, October”, 2000.

[39] J. W. Peirce, “PsychoPy--Psychophysics software in Python”, *J. Neurosci. Methods*, vol. 162, no. 1–2, pp. 8–13, May 2007.

[40] H. Schlosberg, “Three Dimensions of Emotion”, *Psychol. Rev.*, vol. 61, no. 2, pp. 81–88, 1954.

[41] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research”, in *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds. Amsterdam: North Holland Press, 1988.

[42] C. M. Steele, “A threat in the air: How stereotypes shape intellectual identity and performance”, *Am. Psychol.*, vol. 52, no. 6, pp. 613–629, 1997.

[43] J. H. Flowers, “Thirteen Years of Reflection on Auditory Graphing: Promises , Pitfalls , and Potential New Directions”, in *ICAD*, 2005, pp. 1–5.

[44] R. A. Stevenson and T. W. James, “Affective auditory stimuli: Characterization of the International Affective Digitized Sounds (IADS) by discrete emotional categories”, *Behav. Res. Methods*, vol. 40, no. 1, pp. 315–321, Feb. 2008.

[45] C. K. Madsen, “Emotional response to music,” *Psychomusicology*, vol. 16, pp. 59–67, 1997.

[46] C. K. Madsen, “Measurling Musical Response”, *Music Educ. J.*, vol. 77, no. 3, pp. 26–28, 1990.

[47] A.,Väljamäe, T. Steffert, Z. Nagornova, & N. Shemyakina, “Compassion training based on collaborative alpha-asymmetry based neurofeedback”, *Frontiers in Neuroscience*, (in preparation).