A SONIFICATION METHOD FOR MONITORING CHEMICAL SENSOR DATA

Yutian HU[1], **Tony STOCKMAN**[1], **Aleksandar RADU**[2], **Ernesto SAIZ**[3], **and Nick BRYAN-KINNS**[4]

[1]**Queen Mary University of London**, London, United Kingdom
[2]**University of Lincoln** , London, United Kingdom
[3]**Teesside University**, Middlesbrough, United Kingdom
[4]**University of the Arts London**, London, United Kingdom

ABSTRACT

The calibration of chemical sensors is important for ensuring signal integrity. However, for some sensors this can be time-consuming, and without prompt and adequate feedback users may be unaware of calibration errors until post-session analysis is conducted. To address this challenge, we present a real-time sonification framework designed to facilitate efficient monitoring of chemical sensor performance during the calibration stage. The system displays batches of sensor calibration data, collected synchronously, and provides auditory feedback in a sequential time-based manner. To aid users in identifying noisy behaviors in sensor data and comprehending their values, the framework employs three auditory representations: musical note sequences, speech cues, and click sounds. The iterative design process followed a user-centered approach. We detail the iterations of the design, and subsequently evaluate the final approach through a user listening test, discuss benefits and drawbacks of the design, and incorporate user feedback. Our user listening tests conducted using real world data, demonstrate that our method enables efficient detection of abnormal sensor data behaviors.

1. INTRODUCTION

Calibration can be a time-consuming task during the development of novel chemical sensors.

In situations where visual monitoring is restricted or impractical, sonification, the representation of data using sound, provides a solution to the monitoring problem which can be both immediate and intuitive to the user and aid accessibility [1]. In this work, a data sonification system for monitoring the output of chemical sensors without requiring visual interpretation is designed. By transforming certain properties of sensor data into auditory signals, laboratory users are able to gain insights into the condition of sensors by listening to audio without the need for visual representation, thus providing a solution for identifying noise or irregularities in the collected data in real time, that could indicate problematic sensors. The system could also be potentially used to discover data patterns, which may prompt an in-depth analysis on the data at low levels of abstraction.

The primary objective of this project was to explore and develop a sonification system that empowers researchers (users) to easily identify noise or anomalies in data collected during sensor calibration. The user-centered design approach comprised three stages: (1) initial requirement gathering and analysis; (2) iterative design based on expert and user feedback; (3) subjective usability testing.

In the user case in this paper, researchers from University of Lincoln and Teesside University are currently developing a sensor platform which houses up to 16 sensors for continuous monitoring of 4 chemical species in water/soil (4x4 setup, 4 sensors for each analyte). Sensor performance is evaluated based on statistical analysis method, calculation of the slope of the series of data collected from sensors. Data collected from a correctly functioning sensor is expected to have specific patterns. For example, it might show fluctuation at the beginning of each measurement step and it will stabilise towards the end of the measurement.

However, evaluation can be challenging in the early stages under real-world conditions. Sensor signals may deviate unpredictably because of various factors during a calibration session, such as mechanical damage, biofilm formation on surfaces, and the leaching of membrane components. These calibration errors cannot be detected in real-time using current methods, and instead involve processing the data in the laboratory after data collection activities, resulting in a lengthy manual re-calibration process after a certain number of measurements have been taken. In this case, a monitoring method which can help the researchers efficiently detect anomalies during the calibration stage is needed to improve the process and reduce the time spent doing so.

The remainder of this paper is organised as follows: Section 2 provides an overview of sonification methods for monitoring. Section 3 presents the proposed sonification system design process, including the final design, with an emphasis on the user-centered approach applied. Section 4 analyses the results of the user response to a questionnaire for evaluating the sonification system. Section 5 discusses the the main findings of this work. Section 6 concludes the paper by summarizing the findings and the last section listed potential future work. Program code and audio

/ video examples can be found here [1].

2. OVERVIEW OF MONITORING THROUGH SOUND

Monitoring through sound has a long history and is widely in use. The usage of the stethoscope, a medical device for health monitoring, can be dated back to 4000 years ago in the Ebers Papyrus. More recent studies include monitoring a forest through environmental sound recording with sound information retrieval [2] and a security-monitoring system based on audio classification [3]. Though these applications are quite different, they share the same purpose of helping with the detection of unusual states in different conditions through sound. While these cases tend to apply various analysis techniques to get useful information from real-time data, data sonification creates meaning with sound representations of data to convey reliable information.

The International Community for Auditory Display (ICAD) was established 32 years ago with the goal of fostering discussions and shaping the evolving field of sonification. In that time, more sonification techniques have emerged and been developed. Parameter mapping is a very popular technique allowing for high flexibility in continuous mapping between data and different acoustic properties (e.g. timbre, pitch height, loudness). A 2013 study [4] of sonification mapping strategies found that 86.7% of the projects surveyed utilised variable-pitch mapping, rendering it the most popular approach. Such a mapping strategy also appears frequently in monitoring cases. One moderately successful design for monitoring through pitch is the pulse oximeter, a tiny machine broadly used for measuring heart rate [5]. But even with this approach, the study noted that there are a small number of people who cannot identify the pitch change, thus indicating the potential risk in pitch mapping strategies.

Cases in monitoring through sonification apply different mapping strategies and combine sonification techniques to form an efficient data communication. Andrea et al. [6, 7] propose an Electrocardiography (ECG) sonification system that aims to provide users with unobtrusive auditory cues for monitoring their own or their patients' heart signals through water sounds and morphing timbre. Barra, et al. [8] work on sonification of web servers using personalised musical events to describe web status peripherally. Athina and Andreas [9] present a framework for stock market real-time monitoring through Really Simple Syndication (RSS) feed. Concurrent earcons [10] and synthesised speech are used to represent data.

3. DESIGN PROCESS

With the aim of enhancing the calibration process for chemical sensors, we undertook a three stage user-centred design approach [11] to design and test the sonification of chemical sensors during the calibration stage.

[1] Repository for Pure Data patch, audio sample links and introduction video link in this project: https://github.com/Sparasii-Yuna/Sonification-platform-for-chemical-sensors

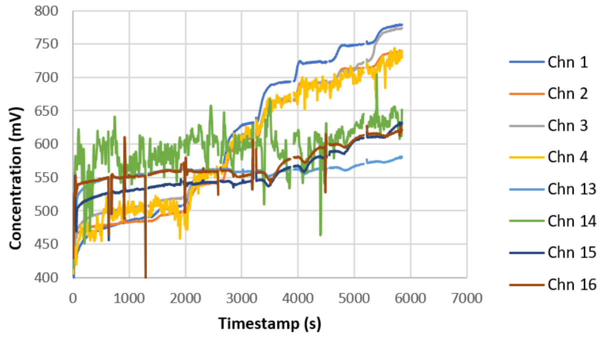

Figure 1. Graph visualisation for example data collected from 8 sensors (Chn1-4 & Chn13-16) during a complete calibration session, combining 9 measurement sessions, with data collected every 8 seconds.

Five participants were involved, including two chemistry researchers with no disabilities (chemistry specialists, also referred to as users in this paper), two experts in sonification, and a sonification designer. Communication between participants included online meetings and emails as well as face to face discussions and focus groups. The design iterations were stored offline in the designer's personal computer and shared throughout the design process. Questionnaires of the final usability listening test are gathered. Audio samples generated in the sonification design were stored on SoundCloud, an online platform for audio content sharing. The three stages of the co-design process are reported in the following sections.

3.1 Stage 1: Requirements Communication

The initial stage consisted of requirements gathering, analysis and early prototyping. The requirements were collected based on comprehensive problem definition documents provided by end users. Additional detail was acquired from verbal explanation on example datasets and research methodologies during two virtual meetings. Example datasets were provided for further analysis.

3.1.1 Current Measurement Precision Estimation Method

Currently there is no real-time analysis by an end user of sensor data - instead, users apply analysis after each calibration session based on statistics and graphs such as in Fig. 1 for sensors performance evaluation.

Details about the statistical analysis method for sensor evaluation are given in [12]. The calculations are derived from the average value of multiple data points collected during each measurement session, where the sensor signal is deemed to be potentially stable. Subsequently, the slope of these averaged values is determined. If this calculated slope falls within the range of what is theoretically considered an appropriate response for the sensor, then the sensor is temporarily classified as reliable. Conversely, if the slope lies outside this range, the sensor's reliability is questioned.

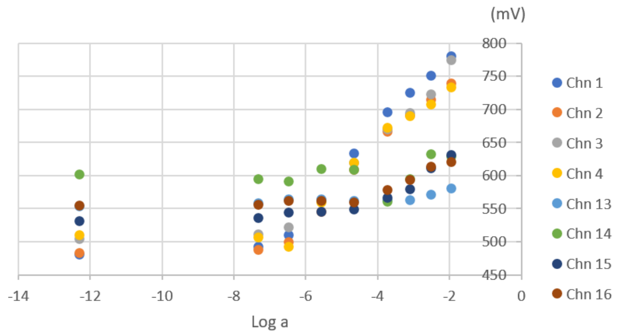

Figure 2. Average concentration values for each of the 8 sensor channels (1-4 and 13-16) across 9 measurement sessions, displayed on a logarithmic time scale.

3.1.2 Dataset

Example sensor data was collected in a laboratory setting. Two different chemical concentrations within a solution were measured with 8 sensors divided into two groups (Group A and Group B), with each group consisting of 4 sensors respectively. On the 4x4 channel sensor platform, Group A was arranged in the first row, occupying slots 1 to 4, while Group B was aligned in the fourth row, taking up slots 13 to 16. For calibration, sensor signals were recorded every 8 seconds over a 10.5-minute period, after which the sensors were removed from the solution. The average of the latest 10 data points which were ideally stablised was calculated for each channel. The most recent data point was excluded due to its tendency to exhibit an abrupt change. This process was repeated across nine measurement sessions, with each session involving an increased concentration level of the target chemical species in the solution. The collected averages from these sessions were then used for the evaluation activities.

Fig. 2 displays all the average values gathered throughout 9 measurement session steps of the dataset. Each slope was calculated through the last 5 points in the graph to ensure that slight deviations at low concentrations are not included. The value of slopes and evaluation results are shown in Table 1. The usability of sensors should be at a good state when the slope value is above 30, theoretically indicating an appropriate sensor response. For example, the data collected from sensor of channel 13 gained a slope value of 4.035, referring to its sensor response being too flat. However, the slope value might fail to reflect the usability. For example, with the slope value of 41.309, sensor channel 14 should be in a good condition. But after observing its signal visualisation in Fig. 1, we could find it extremely noisy, which means the usability is bad and a re-calibration is needed.

3.1.3 Problem Analysis

End users stated that they wish to identify abnormal behaviors in sensor signal as soon as possible during the calibration sessions so that they can take corrective action immediately. Without any real-time monitoring assistance, their current evaluation method relies on calculating the slope

	Chn1	Chn2	Chn3	Chn4	Chn13	Chn14	Chn15	Chn16
Slope	59.104	54.151	55.572	46.400	4.035	41.309	38.273	24.957
Usability (Based on Slope)	good	good	good	good	bad	good	good	bad
Usability (Based on Visualization)	good	good	good	good	bad	bad	good	bad

Table 1. Different usability evaluation results on the same dataset with two methods: slope calculation and visualization observation, showing that the slopes could fail to be used to evaluate the sensors in calibration

of a small number of average numbers from one measurement session, which is highly time consuming. Moreover it is often required to re-calibrate sensors after several measurement steps which compounds the total evaluation effort required. Due to the factors mentioned previously, the sensors are easily affected, thus deterioration in signal is not rare, which makes the situation problematic. Because of this, an efficient monitoring solution is needed to help them evaluate sensor performance with ease and repeatability.

The dataset's structure can be summarized as consisting of 8 parallel abstract one-dimensional series of data. The challenge for sonification lies in analyzing and extracting meaningful variables to represent this data effectively. It is hypothesized that a habitualization process may be necessary for sonification, indicating the need for a systematic approach to familiarize individuals with the auditory representations of the sensor data over time to optimize understanding and interpretation.

Analysis of the problem requirements thus motivates our sonification research in two ways:

- How can we employ sound as an effective means to tell when unacceptable deterioration appears in the sensor signal?
- How can sound inform us of the approximate concentration of the target ion?

3.2 Stage 2: Iterative Design

Iterative design of sonification of sensor data was driven by expert evaluation and user feedback. Design iterations are described in Table 2 and audio samples are available on SoundCloud [2]. During this stage, three online design review meetings were held to gather feedback and ensure the system was meeting the problem requirements. The first meeting evaluated a conceptual design resulting from the previous requirements communication stage and gathering suggestions from experts. The design was developed after the abandoned prototype V0. The design suggested sequential display of data from different channels and mapping data to musical notes instead of direct frequency mapping. A sonification platform was then developed with Pure Data [3] once the specific structure of conceptual design was agreed upon. The second meeting and the third meeting both focused on listening tests and feedback gathering from participants of the meetings. Audio samples were generated of V1 and V2 and shared in advance of

[2] Audio samples of iterations: https://on.soundcloud.com/4xkLX
[3] Pure Data official website: https://puredata.info/

the second meeting for user listening test, while the listening test for V3 took place in the third virtual meeting via a shared screen to share the audio and Pure Data interface simultaneously, so that sound-parameter mappings could be observed in real-time.

Expert evaluation and usability testing with actual users were applied in each prototyping cycle. Experts listened to outputs from our prototype systems and evaluated the quality and usability. For example, during the second meeting, to evaluate the V1 and V2 design, both experts mentioned that the amount of reverb effect in the audio samples of V1 was too much, leading to an optimization for user experience in the next iteration.

3.2.1 Audio Samples

In this section we introduce the audio samples on SoundCloud of different versions of prototypes. The description of representations applied in each iteration is listed in Table 2.

The "V0 FastPrototyping" audio sample, derived from an early prototype during the requirement gathering stage, mapped data values to frequency for simultaneous data presentation. However, due to the challenge of identifying the data source channel in the audio stream, the design was ultimately abandoned.

The audio samples "V1 Chn1-4", "V1 Chn13-16", "V2 Chn1-4", and "V2 Chn13-16" were reviewed in the second meeting. V1 and V2 implemented two variations of design where a fixed set of continuous data was represented as sequences of musical notes following a speech cue, displaying the sequence of four channels individually. It's worth noting that Chn1-4 are for data from well-performing sensors while Chn13-16 are from sensors including problematic ones.

The samples "V3 Chn13-16 Overlapped" and "V3 Chn13-16 NonOverlapping" are examples of the audio generated in the third meeting, where participants observed real-time parameter changes in the system controlled by the designer. the "Overlapped" version had speech cues overlapping with note endings, while the "NonOverlapping" version avoided this overlap.

The final design samples, "V5 AlertRange 10 Chn1-4" and "V5 AlertRange 10 Chn13-16", contrast with "V5 AlertRange 5 Chn1-4" and "V5 AlertRange 5 Chn13-16", which feature a lower alert threshold set at 5. More details will be introduced in the next section.

3.3 Final Sonification Design

The final design employs three types of audio representations as described in this section: musical note sequence (pitch representation), speech cues and click sound.

3.3.1 Musical Note Sequence

The pitch of each musical note is mapped to values of the data, with a higher pitch being related to a larger value, as humans are naturally sensitive to pitch [13], and have the ability to tell whether one tone is higher than another [14]. The motivation is that many listeners should be able to infer a change in value by perceiving difference between two notes, and grasp the data contour through the sequence (melody) contour.

Version	Representations	Description	Feedback
V0	pitch	4 pitch, same timbre, display simultaneously	[abandoned] unable to identify each stream
V1	musical note sequence speech cues	5-note sequence after speech cue, notes with same timbre, short interval between notes, with reverb	too much reverb
V2	musical note sequence speech cues	5-note sequence after speech cue, notes with different timbres (different wavetables), long interval between notes, no reverb	each note can be heard clearly difference between timbres not significant
V3	musical note sequence speech cues	5-note sequence after speech cue, notes with same timbre, short interval between notes, with less reverb, more adjustable intervals	suitable reverb and suitable number of notes need to highlight big changes, ask for no overlaps between speech cue and notes
V4	musical note sequence speech cues click sounds	5-note sequence after speech cue, notes with same timbre, short interval between notes, with less reverb, click sounds with slightly different timbres with granular techniques	[abandoned] click sounds too random
V5 (Final)	musical note sequence speech cues click sounds	5-note sequence after speech cue, notes with same timbre, short interval between notes, with less reverb, click sounds with the same timbre	user feedback (see Section 4)

Table 2. Iterations During Middle Stage Design Process (developed in Pure Data)

We arrange 5 musical notes in a sequence to represent 5 successive data items (from oldest to most recent) collected from one sensor. Each sequence is played with an inter-onset interval (IOI) setting as 150 ms, which could translate to a total sequence duration of 600ms. This timing was chosen to enable listeners to clearly discern the pitch of each individual note, while also ensuring the sequence is brief enough to accommodate the playback of four sequences within an 8-second period — the interval between updates in sensor data. As the dynamic range and value of original data is too large for directly mapping to musical notes, data input is scaled down with an offset to map to notes in a range within the human audition. The scaling strategy also aims to align the pitch variations with dynamic in data. The mission is to keep the pitch at a stable level (with little or no pitch changes) when the data becomes stable, but also to ensure that the musical sound acting unstable (with frequent pitch changes in larger intervals) when the data changes substantially in value. Regarding the timbre of the musical notes, pure tones generated by sine waves with reverb effect are adopted in the final design. Ultimately, a scale factor of 1:25 is applied.

3.3.2 Speech Cues

Four speech cues are used to represent the index of the current sequence about to be played. The speech cues are female speech of four numbers ("one", "two", "three", "four") recorded from the text-to-speech service of Google Translate [4]. With groups of four channels of data to be sonified, speech cues help listeners identify the data source of a sequence. A speech cue is played in advance of the tonal sequence it refers to. The time interval between the onset of a speech cue and its associated sequence is set at 600 ms as this interval is short enough for listeners to group the speech cue with its relevant sequence. It is also long enough to prevent overlap between speech cue and note sequence. However, it was found during the sixth meeting that such overlap can cause annoyance and confusion

[4] https://translate.google.com/

to users. Fig. 3 visualises how the sonification combines speech cues and note sequences in a group of of sensor data channels for five data points per channel.

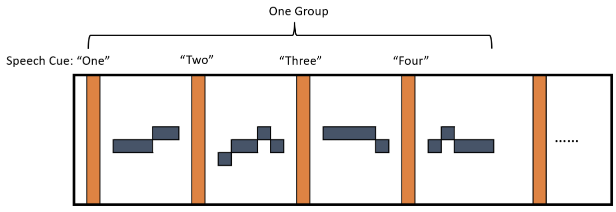

Figure 3. Visual Representation of Speech Cues and Musical Note Sequences

3.3.3 Click Sound

Click sounds function as alert audio signals highlighting abnormal behaviors in the data. As an auditory icon, this representation is intended to be easy for users to learn. An abnormal behavior is defined as the difference between successive data exceeding a fixed threshold (alert range). However, the appearance of abnormal behaviors does not necessarily mean that the sensor is problematic, but a high frequency in abnormal behaviors can imply potential problems. Fig. 4 shows a pair of examples about the mechanism of click sound triggering.

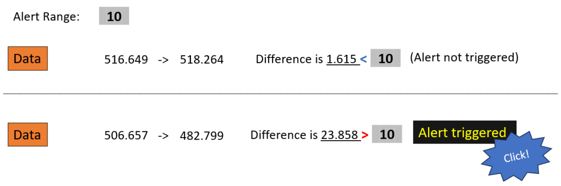

Figure 4. Two click sound mechanism examples. Data is taken from the dataset described in Section 3.1.2.

In the final design the alert threshold is set to 10 based on subjective evaluation of the designer and analysis of 4 example data samples.

Fig. 5 illustrates the frequency of click sounds when the alert range is set to 5 (lefthand of figure) and 10 (righthand of figure). Eight data samples are illustrated in the figures with time axis ascending vertically downwards. An alert range threshold of 5 shows a high frequency of false-positive alerts (regular clicks) in Chn4 (the fourth column from the left), whereas an alert range threshold of 10 does not include these false-positives, suggesting that a threshold of 10 for the alert range reduces the false-positive alerts.

3.4 Stage 3: Evaluating The Final Design

Usability testing of the interactive sonification was undertaken via listening test with a structured questionnaire and an extra interview. As users may be not familiar with the style of listening required for the sonification design, an introductory video was made to explain the mechanics and

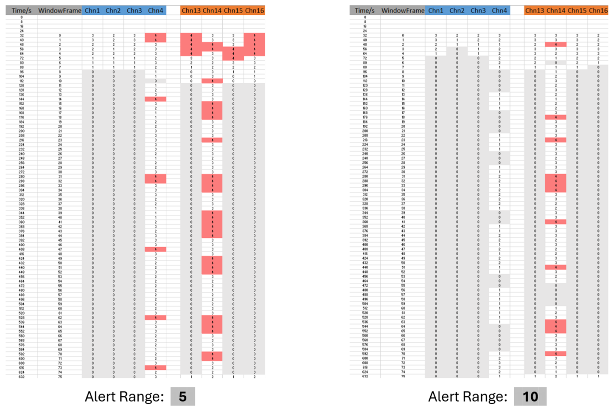

Figure 5. Comparing different alert range threshold settings, 5 (left) and 10 (right). If the number of clicks in a cell is 4 (maximum data points in a sequence), the cell is marked red. If the number is 0, indicating no occurrence in a sequence, the cell is marked as grey.

provide guidance. The video introduced audio representations and their relationship with data. Short audio excerpts that were typical of the system were provided with visualisation and explanations to aid understanding. A questionnaire consisting of 6 questions was administered to the two users to gather qualitative information about usability of the design, particularly in identifying problematic sensors. These questions are listed below:

- *Do you think this design helpful for monitoring data during a calibration session?*
- *Can you observe abnormal behaviours in data streams through this design? If yes, do you think you rely more on pitch information or click sound?*
- *Do you think that the click sound mechanism can accurately reflect the severity of problems in data stream?*
- *Can you easily deduce from this sonification whether there are any problematic sensors within a group and subsequently identify them?*
- *Do you feel it hard to notice any changes in sonified data after listening to it for a period of time?*
- *Any other suggestions/questions/description of feeling?*

4. RESULTS

4.1 Results of Final Stage Evaluation

First we report on the responses to the questionnaires by our two users (U1 and U2) to assess the performance of our sonification system. Both users are familiar with the dataset and the laboratory setting and so provide domain expert feedback on the suitability of the sonification, but they are not experts in sonification and are not musically trained.

Generally the design is regarded as potentially helpful for data monitoring during a calibration session by the users. Both users reported a positive attitude towards the design, evaluating the design as "*in principle helpful*" (U1) and with "*good potential*" (U2).

However, some problems in the design were also identified. One user stated that "*perhaps four sensors are too many*" (U2) which indicates that the information provided in a group exceeded his cognitive load. Beyond that, a correct understanding about the relationship between musical notes and data may have not have been established in users' mind, which may have introduced some misleading interpretations when users were listening to the sample materials. Both users appear to have expected a stable signal to have less variations in pitch, as suggested by the response "*I expect much less variations*" (U2) for stable signal and "*the pitch may be going up or down (due to the jump in value over the threshold) depending on the type of ion*" (U1). These statements may indicate that there is some misunderstanding about the relationship between pitch and data. To be specific, U1 used "*jump in value over the threshold*" to explain the cause to change in pitch / note, which does not match the nature of scaling. With the scaling technique applied, data in specific ranges are mapped to specific notes. For example in our design, with a scaling factor of 1:25, data ranging from 300-325 all map to the musical note G#3. If the data is subtly changing but also changing across the boundary between two specific ranges, the change in pitch would also happen in a sequence. It will act as a "*pitch going up and down*" (U1), which may conflict with users' expectation for a stable pitch in stable signal as their understanding appears to be that change in pitch only happens when jump in value is over certain threshold.

Compared to pitch, click sounds may have improved the performance of the sonification and may be easier for our users to understand. One user "*was able to identify a problematic sensor*" (U2) once they "*were able to focus on it (the click sounds)*" (U2). they "*almost entirely ignored pitch and focused on the click sound*" (U2) and they stated that they are probably "*not able to identify the problematic sensor using the pitch alone*" (U2), which suggests that the click sounds deliver clearer information than pitch to the users. U1 did not express a strong preference for click sounds through their responses, but stated that "*click on a regular basis could mean that the sensor is not working properly (possible failure)*" and "*If only one or two of them are getting the clicks, we may be able to discard its data, or replace them in the field if necessary*". Through these statements, we suggest that U1 understood the representation and was also able to gain more insight into sensor conditions through listening to click sounds.

We found that training is needed in this design to help users familirise themselves with monitoring through sonification. Both users mentioned the importance of training and one user stated that the key is to know "*how it sounds when the sensor is failing and how it is a sensor that it is working*" (U1). We suggest that our introductory video functioned as a useful training guide. Both users relied on the video to make sense of the design. One user watched the video "*many times just to understand the connection between signals in visual terms versus in sound terms*" (U2). The other user commented that the video "*provides the information needed to understand the sonification of the data*" (U1). However, more effort might be needed to explain the relationship between pitch and data clearly. As U2 said, "*I am afraid that I still do not have a full comprehension of how pitch helps*". U2 also suggested "*a slower introduction*" — despite multiple viewings, U2 still felt confused by the musical note sequence mechanism. This implies that the introductory video is helpful but insufficient on its own. To enhance user comprehension, a more detailed and interactive training approach or a longer training session might be beneficial.

4.2 An Extra Interview Session

As some false observations were reported in U1's response, an extra meeting with U1 was held to walk through their responses and discuss the results in more detail. Here the false observation refers the user reporting that they heard click sounds where there were no clicks in the sample audio. The feedback sheet from U1 was also updated after this meeting. During the interview, it was confirmed that the reason for false observations was that user guessed the sonification results with their prior knowledge of the sensors. It was also noted that, besides Channel 14 (Chn14), two additional sensors associated with Channels 13 (Chn13) and 16 (Chn16) were identified as problematic. However comparing to the sensor relating to Chn14 (acting extremely noisy when collecting data), the other two sensors were very stable but did not reflect correct data values as expected. Such failure in sensors cannot be detected by click sounds, as the purpose of click sound representations is to indicate behaviors of large jumps in values of successive data, to inform the users that the signal is noisy. Despite all this, after a discussion with U1, it was shown that this kind of behavior can still be observed through pitch. As U1 stated in the updated feedback: "*After the additional steps during the calibration process two sensors are keeping similar values or going downwards instead of upwards. So, there are no changes in pitch, when there should be.*" Meaning that through pitch representations, users can also potentially identify failing sensors which are not providing responses.

During the meeting, the misunderstanding from U1 about the pitch mechanism was also discussed and corrected. It was found that the users did not manage to acquire enough knowledge about pitch representation through the introduction video, thus leading to an incorrect understanding about the scaling process. U1 suggested adding more details in the instructions to explain the sonification process better, particularly the assignment of each pitch to specific data ranges. After understanding this relationship, U1 realized that pitch information reflects the concentration range of the sample and found the representation more useful. Additionally, U1 concluded that they would rely more on the click sound to identify regular noisy sensors.

5. DISCUSSION

In this section discussions focus on relating the findings in results with the two research questions proposed before.

Our design of click sounds mechanism can effectively inform users of abnormal deterioration in signal. Results of the final usability test shows that both users are able to detect anomaly with click sounds. After a small amount of training, users easily recognised click sounds in an audio stream where musical notes and click sounds are heard simultaneously. What's more, one user reported that only click sounds can be identified after listening for a long time. One possible reason for this might be that the meaning of note sequences was not as clear as click sounds for users. Similar results are also found by Edworthy et al. [15] showing that auditory icon alarms are easier than tonal alarms to be identified either in priority and function domain. As it is easy to adjust parameters (e.g., alert range) in real-time within the Pure Data patch of this work, users could customise the sonification in laboratory settings to tailor it to different needs,

The design of musical note sequence can potentially inform users of the estimated level of target ion's concentration when user has a solid understanding of the mapping strategy. From our results, we suggest that the display of note sequences from 4 sensors in 8 seconds may lead to saturation of cognition load in users quickly, especially with limited training. Compared with the design of the pulse oximeter [5], a successful sonification design which uses pitch to monitoring data from one channel, our sonification displays data from four channels one by one in short delay. The continuity in monitoring one channel is interrupted. Though the trend in data can be observed in the beginning by paying attention to note sequence, users reported to get lost shortly after. It might be because more attention were paid to click sounds during the listening test as the task of identifying noise was more urgent. It was also found that when a tone sequence shares the same contour and a similar tonality with the other sequence just displayed, it will be hard for listeners to discriminate between them [16].

5.1 Future Work

The proposed sonification method in this work may benefit from some additional experiments to fine tune its behavior and improve user satisfaction. Our evaluation was based on the same dataset. Evaluation on different kinds of data inputs would be useful to know and determine if the sonification approach can be generalised to other kinds of data. More experiments need to be carried out to find reliable parameter settings, for example, the setting of the Inter-onset interval (IOI) of note sequences and intervals between groups. More usability tests with a larger focus group with customisable parameters need to be conducted. Finally, experiments investigating concurrent auditory icon and tone sequences would be useful based on this work.

6. CONCLUSION

We suggest that our final sonification design strategy can efficiently help users monitor some abnormal behaviors (noisy activities) in real-time data streaming and identify problematic sensors. Our results support the potential for sonification to aid in these chemical sensors monitoring tasks, and indicate a need to further validate our sonification approach by running them in the real monitoring tasks. During the final usability test, it was found that the click sound representation makes the sonification more useful in identifying the sensors that are acting too noisy, while the pitch representation makes the sonification potentially useful in identifying sensors with inadequate responses. Training sessions may improve the performance of the sonification method by familiarising users with listening modes for monitoring tasks. The introductory video in the final design evaluation was helpful for users to learn how to listen to the sonification but also requires some improvement in explanation of pitch representation. Crucially, however our user-centered design process with participatory design has yielded a successful sonification design with end-user satisfaction. Such a design process can be effective in interdisciplinary scenarios. Above all, our sonification design may potentially be very useful in real-time monitoring task of data collected by chemical sensors.

7. REFERENCES

[1] J. Lazar, S. Chakraborty, D. Carroll, R. Weir, B. Sizemore, and H. Henderson, "Development and evaluation of two prototypes for providing weather map data to blind users through sonification," *J. Usability Studies*, vol. 8, no. 4, p. 93–110, aug 2013.

[2] E. Olteanu, V. Suciu, S. Segarceanu, I. Petre, and A. Scheianu, "Forest monitoring system through sound recognition," in *2018 International Conference on Communications (COMM)*, 2018, pp. 75–80.

[3] A. Abu-El-Quran and R. Goubran, "Security-monitoring using microphone arrays and audio classification," in *2005 IEEE Instrumentation and Measurement Technology Conference Proceedings*, vol. 2, 2005, pp. 1144–1148.

[4] 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.

[5] G. T. Schulte and F. E. Block, "Can people hear the pitch change on a variable-pitch pulse oximeter?" *Journal of clinical monitoring*, vol. 8, pp. 198–200, 1992.

[6] A. L. A. Blanco, S. Grautoff, and T. Hermann, "Cardiosounds: Real-time auditory assistance for supporting cardiac diagnostic and monitoring," in *Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences*, ser. AM '17. New York, NY, USA: Association for Computing Machinery, 2017. [Online]. Available: https://doi.org/10.1145/3123514.3123542

[7] A. L. A. Blanco, S. Grautoff, and T. Hermann, "Heart alert: Ecg sonification for supporting the detection and diagnosis of st segment deviations," in *5th Interactive Sonification Workshop (ISon 2016)*, 2016, pp. 1–8.

[8] M. B. T. C. A. De, S. U. F. Petrillo, A. Negro, V. Scarano, T. Matlock, and P. P. Maglio, "Personal webmelody: Customized sonification of web servers," 2001.

[9] A. Bikaki and A. Floros, "An rss-feed auditory aggregator using earcons," in *Proceedings of the 6th Audio Mostly Conference: A Conference on Interaction with Sound*, ser. AM '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 95–100. [Online]. Available: https://doi.org/10.1145/2095667.2095681

[10] S. A. Brewster, P. C. Wright, and A. D. N. Edwards, "Experimentally derived guidelines for the creation of earcons," 2001. [Online]. Available: https://api.semanticscholar.org/CorpusID:17462746

[11] T. Robertson and J. Simonsen, "Participatory design," *Routledge international handbook of participatory design*, p. 1, 2012.

[12] P. W. Dillingham, T. Radu, D. Diamond, A. Radu, and C. M. McGraw, "Bayesian methods for ion selective electrodes," *Electroanalysis*, vol. 24, no. 2, pp. 316–324, 2012.

[13] L. Demany and F. Armand, "The perceptual reality of tone chroma in early infancy," *The journal of the Acoustical Society of America*, vol. 76, no. 1, pp. 57–66, 1984.

[14] J. Plantinga and L. J. Trainor, "Memory for melody: Infants use a relative pitch code," *Cognition*, vol. 98, no. 1, pp. 1–11, 2005.

[15] J. R. Edworthy, C. J. Parker, and E. V. Martin, "Discriminating between simultaneous audible alarms is easier with auditory icons," *Applied Ergonomics*, vol. 99, p. 103609, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0003687021002568

[16] W. J. Dowling, "Tonal strength and melody recognition after long and short delays," *Perception & Psychophysics*, vol. 50, pp. 305–313, 1991.