Design System Analysis to Create User-Interface Guideline for a Sound Synthesis-based Audio Application

Shareen Rhema Angela Gumulya [1], Arya Harditya[2], Tabitha Mira Emanuelle[3], Tombak Matahari[4], Fransiskus Rendy[5]
[1]Sampoerna University, Indonesia [2]Sampoerna University, Indonesia [3]Akademi Komunikasi SAE Indonesia, Indonesia [4]Sampoerna University, Indonesia, [5]Akademi Komunikasi SAE Indonesia, Indonesia
shareen.gumulya@my.sampoernauniversity.ac.id
arya.harditya@sampoernauniversity.ac.id
tm.emanuelle4@gmail.com
tombak.matahari@sampoernauniversity.ac.id
r.retanubun@sae.edu

Abstract. Sound design is a critical aspect of audio engineering, especially in the field of sound reproduction. There are two commonly used methods for creating new sounds: sound synthesis and sample recording. Sound synthesis is the process of creating sound from scratch using either hardware or software tools. The tool presented in this research is a sound synthesizer that was created using the Max 8 application. It offers a more unique user experience compared to commonly used sound synthesizers. Before conducting this research, we had developed a tool for sound design based on sound synthesis. During the production of this sound synthesizer, one of the main challenges was developing the user interface. Arguably, this may be attributed to audio engineers' tendency to prioritize sonification over visualization. The analysis of UI design guidelines comprises five components of usability quality: learnability, efficiency, memorability, errors, and user satisfaction. The final product undergoes testing and evaluation using post-task questionnaires, specifically the System Usability Scale (SUS). This article explores the design of the user interface (UI) for an experimental sound synthesizer. This is significant for the audio engineering community because it sets usability standards for the design of sound creation tools in the future.

Keywords: Sound design, user interface, sound synthesis, system usability, design

1 Introduction

Sound design is a critical aspect of audio engineering, especially in the field of sound reproduction. Another definition of sound design is the process of working with the sound of an object in the early stages of sound development to convey desired information or feelings [1]. Horror movies use sound design to evoke fear and anxiety based on human psychology. Whittington's proposal highlights that both horror and sound design aim to immerse viewers in scenes of anxiety, terror, and dread, tapping into unconscious fears. The unknown and mysterious elements in horror films trigger the subconscious fear of death, strange phenomena, and unfamiliar spaces. These non-existent or undiscovered scenes challenge normal cognition, intensifying the sense of

© The Author(s) 2024
T. Ratuannisa et al. (eds.), *Proceedings of the 3rd International Conference of Art, Craft, Culture and Design (ICON-ARCCADE 2023)*, Advances in Social Science, Education and Humanities Research 841,
https://doi.org/10.2991/978-2-38476-238-5_10

mystery. Similarly, unconventional sounds in horror movies contribute to this sense of strangeness, disrupting viewers' perceptions and evoking instinctive fear [13].

The subjective nature of sound is a crucial aspect of horror movie soundscapes. This technique effectively places the audience in the emotional state of the protagonist, where even minor sudden sounds demand strong attention due to human instinct. This instinct dates back to ancient times when creating noise was believed to ward off evil and boost courage, exemplified by practices like Chinese firecrackers. The sudden introduction of silence or sound variations in familiar settings triggers psychological responses, as demonstrated in films like "The Other," where eerie floor creaks and ethereal piano notes induce an atmosphere of terror without explicit visuals. These cinematic symbols convey impending crises and horror effectively. Another example is the simple TV noise in "Ring" accompanying a sinister videotape, enhancing anxiety and intensifying the sense of oppression surrounding the character Sadako. This manipulation of sound immerses the audience, blurring their connection to the actual environment. These sounds provide crucial story development and encourage audience imagination, ultimately maximizing the horror atmosphere. The earliest sound films primarily focused on synchronizing dialogue, causing filmmakers like Sergei Eisenstein and his associates to express reservations about whether "talking films" were simply imitating dramatic art. This led them to propose the concept of "contrapuntal aurality," laying the foundation for a sound design technique that eventually found its place in horror films [13]. Contrapuntal sound, born from this theory, involves the simultaneous coexistence of visual and auditory elements in a film. These elements follow distinct narrative paths, yet they are ingeniously combined based on their trajectories. This fusion creates a holistic effect that surpasses the impact of isolated visuals or sounds. The structure of sound and picture counterpoint represents an evolution in the interplay between audio and visual components. It allows sound and images to transcend mutual dependence and repetition, enabling them to fulfil their distinct roles. This technique contributes to character development, metaphorical expression, and the unique enhancement of the film's atmosphere.

There are two commonly used methods for creating new sounds: sound synthesis and sample recording. Sound synthesis is the process of creating sound from scratch using either hardware or software tools called synthesizers. The user interacts with the synthesizer to create sounds at the most basic level. This frequently entails using a keyboard or other device to directly manipulate the synthesizer, this is referred to as a "direct manipulation" system, which has proven to be significantly more usable by novices [10]. Thus, the synthesist is acting in a capacity similar to that of an orchestral player. However, when the control is indirect, such as when using MIDI or other computer-based methods to remotely control a synthesizer, the position is much more akin to that of an orchestra conductor with the option of also acting as a solo performer. In a conventional instrument, the control interface is often predetermined by the instrument itself. For example, a guitar has six strings, a fretboard, a bridge and so on and the player can pluck, strum, or tap the strings, which may be open or fretted. A synthesizer, however, does not have a set of fixed controllers because it is not limited by any physical properties. The adaptability of the synthesis technique and its practical implementation, then, decide the control [9]. A musical performance is an amalgamation of an instrument and a player. The interaction between the two produces the music, and the interfacing between the player and the instrument affects the control of the instrument [9]. Interfaces in audio technology have been previously discussed, specifically in a Digital Audio Workstation (DAW) which is primarily a visual environment graphically represented on a computer screen. Since the 1980s, graphic user interfaces have been a crucial component of software development and are meant to provide users with the most natural and uncomplicated way to get the functionality they need [6]. Apart from DAW, Max 8 is also categorized as a music sequencing tool referred to as a music visual programming tool and this system are graphical by definition, as user build their composition system, they can see the “objects”, their parameters’ values, and the links between them. The tool presented in this research is a sound synthesizer called Monsterizer or MNZTR, which was created using the Max 8 application and uses frequency modulation (FM) and wave shaping, a non-linear function transformation of simple waveforms to create complex spectra [11]. It offers a more unique user experience compared to commonly used sound synthesizers. The purpose of this tool is to enhance the artist's or sound designer's ability to express themselves using the software.

Sound in multimedia, movies, games, virtual reality, and human-computer interfaces is a growing field that encompasses the disciplines of analogue and digital signal processing, physics, speech, music, perception, and computer systems architecture [11]. MNZTR is intended for use by sound designers or sound effects editors in audio post-production for the horror or thriller genres. These genres pose a unique challenge because they require rich, dynamic, and evocative sound design. On the other hand, to provide such an experience, the interface must be visually appealing, user-friendly, and easy to comprehend, ensuring both practicality and expressiveness. In conceiving new sequencing applications such as MNZTR, there is a need for a high-level analysis examining the major characteristics of such applications and how these affect their usability [7]. As Max 8 lacks interface-related documentation, the designers in this research project are tasked with analyzing the design system of Max 8 and using it as a guide for developing the UI. During the production of MNZTR, one of the main challenges for the audio engineers was designing the interface. Arguably, this may be attributed to audio engineers' tendency to prioritize sonification over visualization. Similarly, many User-Experience (UX) designers’ backgrounds, schooling and professional training were in the areas of visual design not in audio or sound design [8]. However, audio and visual relationships have been explored in the field of Human-Computer Interaction (HCI) to enhance the user experience and usability, across different application areas – such as accessibility in assistive displays [12]. Therefore, effective solutions require cross-disciplinary collaboration between audio engineers and UI designers. Past research has discussed the relationship between sound and brand identity, although it only mentioned that jingles, sound logos and non-musical sounds are also elements that construct brand identity [1]. While brand identity may be discussed, this article strictly focuses on measuring the system usability of MNZTR, specifically UX and UI. Finally, this article will measure the usability improvement of a sound synthesizer using the System Usability Scale (SUS) in the preliminary test and post-test comparative method.

2 Methodology

2.1 A Subsection Sample

Generally, the usability of any tool or system has to be viewed in terms of the context in which the tool or system is used, and whether it is appropriate to that context [4]. Usability refers to the quality of a user's experience when interacting with products or systems, including websites, software, devices, or applications [2]. Thus, the sound synthesizer application in this article is suitable for the analysis of UI design guidelines which comprises five components of usability quality: learnability, efficiency, memorability, errors, and user satisfaction [3]. As mentioned, the lack of appropriate documentation has hindered the development of this sound synthesizer's UI. Consequently, the synthesizer has not undergone user testing yet, and it is impossible to specify system usability without initially defining the intended system users, how the users perform tasks with the system, and the physical characteristics, and organizational and social environment in which the system is used [4].

2.2 System Usability Scale (SUS)

The final product undergoes testing and evaluation using post-task questionnaires, specifically the System Usability Scale (SUS). SUS is a simple, ten-item Likert scale giving a global view of subjective assessments of usability that a statement is made, and respondents indicate the degree of agreement and disagreement [4]. After literature research, this is possibly not the first time that a sound synthesizer is using SUS to measure its usability, and it should become the standard [10]. As seen in Table 1, the goal of conducting SUS is to receive feedback that can be used to improve the perceived usability of the entire sound synthesizer system. This is significant for the audio engineering community because it sets usability standards for the design of sound creation tools in the future.

Table 1. System Usability Scale by Digital Equipment Corporation in 1986

	Strongly Disagree				Strongly Agree
1. I think that I would like to use this system frequently	1	2	3	4	5
2. I found the system unnecessarily complex	1	2	3	4	5
3. I thought the system was easy to use	1	2	3	4	5

4. I think that I would need the support of a technical person to be able to use this system	1	2	3	4	5
5. I found the various functions in this system were well-integrated	1	2	3	4	5
6. I thought there was too much inconsistency in this system	1	2	3	4	5
7. I would imagine that most people would learn to use this system very quickly	1	2	3	4	5
8. I found the system very cumbersome to use	1	2	3	4	5
9. I felt very confident using the system	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system	1	2	3	4	5

As previously mentioned, usability is measured through its 5 quality components, and to achieve an effective score and testing feedback, the SUS statements in this article are mapped according to the components, as seen in Table 2 below. The purpose of this mapping is to reach a consistent usability outcome based on both the quality and the system; SUS alone was only intended to measure perceived ease-of-use (single dimension). However, more recent research in 2009 shows that SUS provides a global measure of system satisfaction and sub-scales of usability and learnability where statement 4 and 10 represents the Learnability dimension, and the remaining statements provide the Usability dimension [5]. Thus, it has a linear focus conclusion rather than misperception.

Table 2. SUS statements mapped to the 5 Usability quality components.

Quality Components	SUS Statements
Learnability	I think that I would need the support of a technical person to be able to use this system (4). I needed to learn a lot of things before I could get going with this system (10).
Efficiency	I thought the system was easy to use (3). I would imagine that most people would learn to use this system very quickly (7).

Memorability	I found the system unnecessarily complex (2). I found the system very cumbersome to use (8).
Errors	I found the various functions in this system were well integrated (5). I thought there was too much inconsistency in this system (6).
Satisfaction	I think that I would like to use this system frequently (1). I felt very confident using the system (9).

2.3 Usability Score

The average SUS score from all 500 studies is 68 and it is not percentages. A SUS score above 80 is an "A", and score mean of 68 is considered a "C" and anything below 51 is an "F". Subtract 1 score for each odd statement's response, i.e., n-1 (n is the response), and subtract 5 with each even statements' response i.e., 5-n. Sum all the converted responses of each user and multiply by 2.5, which will convert the range of possible values from 0 to 100 instead of 0 to 40 [5].

1. Result and Discussion

The design decisions made for *MNZTR* are mainly based on the brand values that have been decided since the beginning of the project and adjusted further before the redesign. Firstly, the idea possesses practicality, making it applicable to real-world projects and at the same time providing space for innovation. Its defining characteristic and appeal lie in its distinct functionality, a sound synthesizer catering specifically to the horror and thriller genre. Finally, accessibility is ensured through its user-friendly interface, allowing smooth usability to a wider spectrum of users. Through this trifecta of practicality, characteristic functionality, and accessibility, the new user interface and experience of *MNZTR* are executed to reflect these brand values. As seen in Figure 1, the logo is also redesigned to reflect these points, encompassed as a sawtooth wave, the kind of sound wave generated by *MNZTR*.

Fig 1. Logo redesign with a clearer relationship to the characteristic sound produced by the sound synthesizer.

As seen in Figure 2, (a) is the previous UI design and (b) is the redesigned version. The UI redesign is based on one of the design principles, which is the design hierarchy. The power button is on the top-left and the main synthesis process is a larger XY space compared to the previous version. All the buttons are rectangular-based buttons, which makes it easier for users to identify their tactile functionality, i.e., rectangular is to press, and circular is to rotate. The previous version's usability was considered as a pre-test and the redesigned was the post-test.

Fig 2. (a) Previous User-Interface of MNZTR for Pre-Usability Test; (b) Redesigned User-Interface for Post-Usability Test.

The main color green is taken from the previous design as it represents the brand and its function accordingly. The red colours in the previous UI are removed to improve readability. The neon green colour is used to show that the buttons, knobs, and pointer in the XY space are actively being used. In its inactive mode, the controls are a light grey colour. The new buttons and knobs are also inspired by controls that are normally found in audio engineering applications and tools like midi controllers, to create a familiar experience for audio engineers and sound designers.

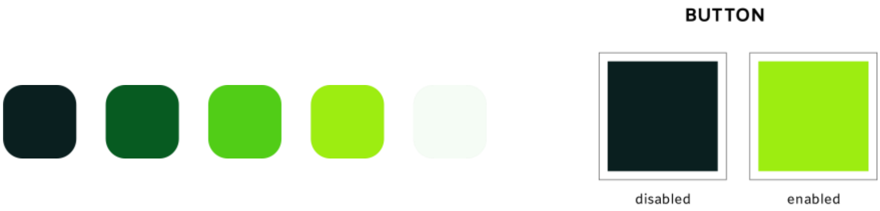

Fig 3. Applying brand colours as UX elements, on and off buttons.

The two results have undergone preliminary and post-SUS tests on 7 users. The calculated results are as seen in Table 3 and both scores are on D grade or Poor usability score. There is no significant score difference between preliminary and post, however, redesigning the UI improves the score but it is unsuccessful in reaching the usability standard. Unfortunately, the result does not show a clearer understanding of which part of the UI needs some more improvements.

Table 3. Pre and post-SUS final score.

Participants	Pre-Test Score	Post-Test Score
User 1	33	30
User 2	78	78
User 3	45	68
User 4	73	53
User 5	75	95
User 6	53	60
User 7	85	83
Average Score	**63**	**66**

As seen in Figure 3, these are the results for the 5 usability components score. There are fluctuations in the pre and post-test components' scores, however, 60% of the components are improved. Efficiency and Errors need to be re-evaluated in future research as both had significant decreases after redesigning. Once again, the result does not show a clearer conclusion of why there is a decrease in both components. There are several potential factors, first, colour harmony in design is not correlated to efficient usability. Some contrast in the colour of buttons and labels could help in improving this. Second, the design pattern is not like any other software that the user could be familiar with. The biggest difference that the *MNZTR* has is the XY space in the middle, which is the main control of the system. This might confuse and even lead to the perception of inconsistency because it deviates too much from commonly known designs.

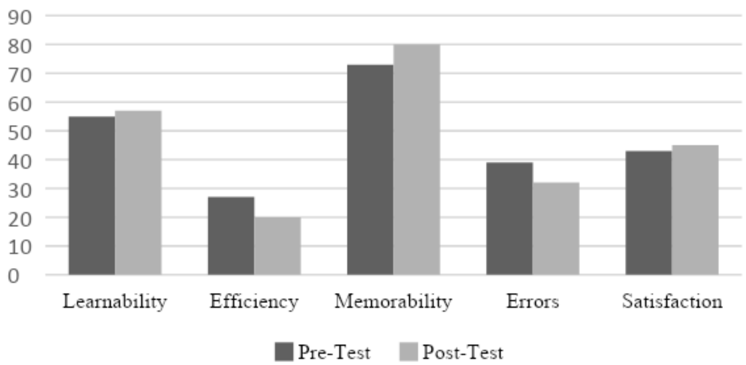

Fig 3. 5 Usability components' score.

Thirdly, redesigning a simpler interface can also be perceived as inconsistent with the characteristic function of the software. The previous design had buttons that resembled slime and represented the horror genre well, whereas the new design offers a sleeker and more modern look, removing some of that character.

In conclusion, getting used to a new design also contributes to the user's workflow and can hinder their ability to complete tasks seamlessly. The use of the joystick to control the software is also a contributing factor to inconsistencies in the system, as the interface does not resemble said joystick. Further user testing must be done to identify inconsistencies and make the necessary adjustments to fit the user's needs and solve their problems. Furthermore, qualitative data may need to be collected for further studies, especially to investigate the visual usability of the synthesizer. Designers need to understand why the new UI is not successful enough to meet the usability standard. In future research, designers should be part of the synthesizer creation process instead of being involved only to improve on what has already been done and decided.

References

1. Carron, M., Dubois, F., Misdariis, N., Talotte, C., & Susini, P. (2014). Designing sound identity. Proceedings of the 9th Audio Mostly: A Conference on Interaction With Sound. https://doi.org/10.1145/2636879.2636896 American Psychological Association. (2020). Publication manual of the American Psychological Association (7th ed.).
2. Department of Health and Human Services. (2013, October 8). Usability evaluation basics. Usability.gov. https://www.usability.gov/what-and-why/usability-evaluation.html

3. Nielsen, J. (2012, January 3). Usability 101: Introduction to usability. Nielsen Norman Group. https://www.nngroup.com/articles/usability-101-introduction-to-usability/
4. Brooke, J. (1995). SUS: A “quick and dirty” usability scale. Usability Evaluation In Industry, 207–212. https://doi.org/10.1201/9781498710411-35
5. Sauro, J. (2011, February 3). Measuring usability with the system usability scale (SUS). MeasuringU. https://measuringu.com/sus/
6. Marrington, M. (2016). Innovation in Music 2015 (InMusic’15). In Innovation in Music II (pp. 52–63). Cambridge; Future Technology Press. Retrieved August 7, 2023, from https://www.academia.edu/29642900/Paradigms_of_Music_Software_Interface_Design_and_Musical_Creativity.
7. Duignan, M., Noble, J., & Biddle, R. (2005, January). A Taxonomy of Sequencer User-Interfaces. Retrieved August 7, 2023, from https://www.researchgate.net/publication/228363177_A_taxonomy_of_sequencer_user-interfaces
8. Follett, J. (2007, June 18). Audio and the User Experience. UXmatters. https://www.uxmatters.com/mt/archives/2007/06/audio-and-the-user-experience.php
9. Russ, M. (2019). Applications: Sound-making Techniques & Controllers. In Sound Synthesis and Sampling (3rd ed., pp. 415–505). essay, Routledge.
10. Scarlatos, A. (2020). Sonispace: A Simulated-Space Interface for Sound Design and Experimentation. https://doi.org/10.48550/arXiv.2009.14268
11. Cook, P. R. (2002). Sound production and modeling. IEEE Computer Graphics and Applications, 22(4), 23–27. https://doi.org/10.1109/mcg.2002.1016695
12. Correia, N., & Tanaka, A. (2021). From GUI to avui: Situating audiovisual user interfaces within human-computer interaction and related fields. EAI Endorsed Transactions on Creative Technologies, 8(27), 169913. https://doi.org/10.4108/eai.12-5-2021.169913
13. Gong, Z., & Zhang, J.. (2021). Review of Sound in Horror Movie. 767–771. https://doi.org/10.2991/assehr.k.211120.139

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.