ASMRcade: Interactive Audio Triggers for an Autonomous Sensory Meridian Response

Silvan Mertes
silvan.mertes@informatik.uni-augsburg.de
Augsburg University
Human-Centered Artificial Intelligence
Augsburg, Germany

Marcel Strobl
marcel.strobl@student.uni-augsburg.de
Augsburg University
Human-Centered Artificial Intelligence
Augsburg, Germany

Ruben Schlagowski
ruben.schlagowski@informatik.uni-augsburg.de
Augsburg University
Human-Centered Artificial Intelligence
Augsburg, Germany

Elisabeth André
andre@informatik.uni-augsburg.de
Augsburg University
Human-Centered Artificial Intelligence
Augsburg, Germany

ABSTRACT

Autonomous Sensory Meridian Response (ASMR) is a sensory phenomenon involving pleasurable tingling sensations in response to stimuli such as whispering, tapping, and hair brushing. It is increasingly used to promote health and well-being, help with sleep, and reduce stress and anxiety. ASMR triggers are both highly individual and of great variety. Consequently, finding or identifying suitable ASMR content, e.g., by searching online platforms, can take time and effort. This work addresses this challenge by introducing a novel interactive approach for users to generate personalized ASMR sounds. The presented system utilizes a generative adversarial network (GAN) for sound generation and a graphical user interface (GUI) for user control. Our system allows users to create and manipulate audio samples by interacting with a visual representation of the GAN's latent input vector. Further, we present the results of a first user study which indicates that our approach is suitable for triggering ASMR experiences.

CCS CONCEPTS

• **Human-centered computing** → **Interactive systems and tools**; **User studies**; **Auditory feedback**; **Graphical user interfaces**.

KEYWORDS

ASMR; Interactive ASMR; ASMR Generation; GAN; Stress; Audio Synthesis

ACM Reference Format:
Silvan Mertes, Marcel Strobl, Ruben Schlagowski, and Elisabeth André. 2023. ASMRcade: Interactive Audio Triggers for an Autonomous Sensory Meridian Response. In *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI '23), October 09–13, 2023, Paris, France.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3577190.3614155

ICMI '23, October 09–13, 2023, Paris, France

ACM ISBN 979-8-4007-0055-2/23/10.
https://doi.org/10.1145/3577190.3614155

1 INTRODUCTION

In today's fast-paced world, stress is increasingly becoming a pervasive and challenging problem in people's daily lives [44, 48]. Persistent stress can lead to a variety of physical and mental health issues, such as anxiety, depression, sleep disorders, heart disease, high blood pressure, headaches, and many other health problems [13, 16, 19, 25, 32, 45]. Therefore, it is very important for people to find an adequate balance to relax and thus counteract stress.

An effective method of relaxation that is becoming more and more popular is the consumption of content triggering an *Autonomous Sensory Meridian Response* (ASMR) [1]. ASMR is a term used to explain the tingling sensation some people experience in response to certain, mostly auditive, stimuli. Typically, this feeling is described as a delightful buzz that begins in the scalp and travels down the neck and back [2]. Auditive triggers for ASMR can vary, including sounds such as whispering, soft speaking, tapping, or scratching. Additionally, visual or physical stimuli, like observing someone do a repetitive task or receiving a gentle touch, can also elicit ASMR. In recent years, ASMR has become increasingly popular, with numerous content creators producing videos and audio recordings intended to induce ASMR in their audience. While there is limited scientific research on the topic, many individuals find ASMR to be a relaxing and pleasurable experience [1]. However, it's important to note that not everyone experiences ASMR or is affected by sounds meant to induce ASMR in the same way.

ASMR is often considered a *passive* experience because it involves a state of relaxation and receptivity to certain stimuli rather than an active engagement with them. However, Barret et al. [1] found a correlation between the number of ASMR responses and the *flow* state of participants, which is the psychological state of being immersed in a task [5]. Hence, as such tasks typically involve active participation or interaction with the environment, e.g., in sports or playing video games, we propose an *active* approach for ASMR triggers that interactively engages ASMR recipients.

To this end, this paper introduces *ASMRcade*, an interactive application for ASMR sound synthesis, exploration, and customization. The application utilizes a Generative Adversarial Network (GAN) that we trained to generate high-quality audio samples that are aimed to trigger ASMR. By providing a graphical user interface, users of ASMRcade are able to interact with that model's latent space and, as such, directly influence the sounds that the GAN produces in real time.

The main contributions of this paper are as follows:

- We trained a WaveGAN [8] model on a dataset of ASMR sounds. Specifically, for a proof-of-concept, we use *tapping* sounds, which have proven to induce ASMR for many people in the past [2]. Such sounds are typically created by tapping on various objects, such as surfaces, or directly on a microphone.
- Using a web-based graphical user interface as introduced by Schlagowski et al. [37], we made the input latent space of the trained WaveGAN model interactive. Specifically, we represent parts of the latent model space of the WaveGAN through certain interface elements. By moving those elements, the user directly influences the sounds that the model produces.
- We present a first exploratory user study in which we let users interact with our system. We show that our system can induce ASMR and, as such, is a promising direction for future research in the field of ASMR.

To the best of the authors' knowledge, this paper is the first work introducing an interactive approach to the field of ASMR research. The online demonstrator using this GUI and the GAN described above can be found online on our web server for demonstration purposes[1].

[1]www.hcai.eu/asmrcade

2 RELATED WORK

2.1 Autonomous Sensory Meridian Response

Autonomous Sensory Meridian Response (ASMR) is a sensory phenomenon characterized by a pleasurable tingling sensation with its origin in the scalp, moving down the neck, and sometimes following the line of the spine down to other areas of the body in response to specific audio and visual stimuli. In recent years, there has been a growing interest in understanding the triggers that lead to ASMR experiences and the potential benefits of experiencing ASMR. As such, the research community has begun to embrace the potential of ASMR as a valuable source of inspiration for Human-Computer Interaction [20].

Barratt et al. [1] proposed that ASMR is a flow-like mental state, which refers to a state of intense focus and absorption in an activity. They suggested that ASMR experiences involve a deep immersion in the sensory stimuli, which can lead to a sense of timelessness and an altered state of consciousness. The researchers also found that people who experience ASMR report a range of positive effects, such as relaxation, stress relief, and improved mood.

In their later study, Baratt et al. [2] investigated the sensory determinants of ASMR by surveying a large online community of individuals who experience ASMR. They found that the most common triggers were related to auditory stimuli, such as whispering, tapping, and scratching sounds. Other common triggers included visual stimuli, such as hand movements and personal attention, and touch-related stimuli, such as gentle touch and hair brushing. Interestingly, the researchers also found that individual factors, such as personality traits and mood, could influence the likelihood of experiencing ASMR.

Poerio et al. [33] further explored the triggers of ASMR by conducting laboratory experiments in which participants listened to a range of audio and visual stimuli while their physiological responses were measured. They found that ASMR experiences were associated with a reduction in heart rate and an increase in skin conductance responses, indicating a relaxation response. The study also found that the most effective ASMR triggers were those that involved interpersonal closeness, such as whispering and soft-spoken voices.

In terms of content, people consume a variety of ASMR videos and audio recordings to experience the tingling sensation. Common types of content include role-playing scenarios, where the so-called *ASMRtist* portrays a particular role (such as a hairdresser or a doctor), soundscapes, where various sounds are played, and guided meditations or affirmations. There are many successful "ASMRtists" with millions of subscribers on YouTube, indicating the existence of a vast online community interested in the phenomenon.

Overall, ASMR is a complex phenomenon that involves a combination of sensory and personal factors, and triggers can vary widely between individuals. However, research has shown that certain types of stimuli, such as gentle sounds and interpersonal closeness, are more likely to elicit an ASMR response. Furthermore, the flow-like mental state experienced during ASMR has been proposed to have potential benefits for mental health and well-being.

2.2 Customizable Sound Generation

The technical foundation to make an interactive ASMR experience possible is a method to synthesize artificial sounds in a controlled way. While early works in that field made use of procedural approaches like *Concatenative Sound Synthesis* [40] or *Modular Sound Synthesis* [46], mechanisms using machine learning are on the rise since several years. For example, approaches from the field of Evolutionary Computing are widely used for optimizing sound synthesis problems [22, 27, 28] and were, although not in an interactive setting, also applied to ASMR sound synthesis [30]. One promising approach for sound generation is the use of Generative Adversarial Networks (GANs), which were originally developed for image synthesis [14]. Various modifications to the original GAN architecture were presented that address the generation of audio data, such as *GANSynth* [11] or WaveGAN [9].

In order to add the possibility for controlling the audio outputs generated by a GAN, mechanisms were proposed that incorporate directive features into the GAN training [10, 24], or directly search through the input space of an already trained GAN, e.g., using methods such as Latent Variable Evolution (LVE) [26]. In the context of ASMR, Fang et al. [12] adapted the DCGAN architecture [34] to create random new ASMR sounds. A similar approach was followed by Oh et al. [31], who created artificial ASMR sounds using the SpecVQGAN architecture [17]. However, although evaluations of both of these approaches indicated that GANs could create auditively pleasing ASMR sounds, they did not include mechanisms to engage with the sound synthesis process interactively.

2.3 Interactive Approaches to Parameter Space Exploration

The generation of ASMR audio content is only one part of our objective - another goal is to actively engage users in the ASMR sound generation process. For the ASMRcade application, we use the WaveGAN architecture for sound synthesis, which can transform an arbitrary, non-interpretable input vector to sound that resembles the data the model was trained on. Therefore, we have to enable the user to interact with the input parameter space of the WaveGAN. In the general field of audio synthesis, multiple approaches exist to interactively explore parameter spaces, such as mapping parameters to 2D-Interfaces using Hilbert curves [47], or using interactive evolutionary algorithms [6, 36]. In the context of music synthesis, interaction with parameter spaces was even used as musical instruments [3, 18, 29, 42].

Further, several works have focused on building interactive systems specifically for exploring a GAN's input parameter space. Here, a popular use case is to interactively control the output of GANs that were trained to generate drum sounds. For instance, Ramires et al. [35] used *SeFa*, a closed-form factorization method [41], to find dimensionality-reduced directions of a GAN's parameter space, which they made accessible to the user through a Graphical User Interface (GUI). Further, Schlagowski et al. [37, 39] aimed for interactive drum sound synthesis for a drum sequencer application. Similar to our work, they used the WaveGAN architecture and built a user interface to directly modify single latent vector elements of a GAN [37, 39]. As they observed positive user experience ratings in their user study, we decided to include certain GUI-elements of their system within our system.

3 THE ASMRCADE SYSTEM

Our system consists of two major parts. Firstly, a WaveGAN model that was trained to produce highly realistic and diverse ASMR tapping sounds, and secondly, a web-based graphical user interface, which includes visual representations of the WaveGAN's latent input vector elements. By interacting with those visual representations, i.e., changing their spatial position, the user can directly influence the WaveGAN's audio output in real-time. All in all, by using our system, the user can actively engage in the ASMR sound generation process instead of just passively listening to ASMR stimuli. The following sections give a more detailed overview of the ASMRcade system's single components.

3.1 Audio Generation

In recent years, *Generative Adversarial Networks* (GANs) have excelled in creating high-quality, diverse data that is nearly indistinguishable from data that was used for training ("real" data). GANs were first introduced by Goodfellow et al. [15] and consist of two neural networks - a *generator* and a *discriminator* - which compete against each other in a min-max game. During training, the generator learns to convert random noise vectors to new, artificial data, while the discriminator learns to distinguish between that data and real data from a training dataset. As such, the generator tries to fool the discriminator while the discriminator tries not to be fooled by the generator. A modification of the original GAN framework was introduced by Donahue et al. [8], who presented the WaveGAN architecture. While originally, GAN models were used for image synthesis, WaveGAN was specifically designed for audio synthesis. We selected the WaveGAN architecture for our system as it is fast and fulfills the required quality standards of the generated output for successful ASMR triggers.

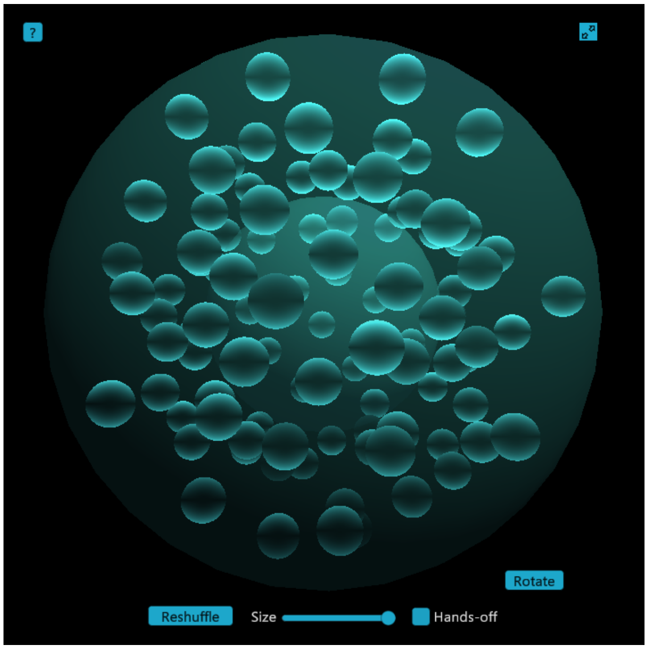

Figure 1: The ASMRcade user interface adopted from Schlagowski et al. [37].

3.1.1 Dataset. A necessary prerequisite for training a WaveGAN model is the existence of a suitable dataset. Datasets that can be used to train a GAN for high-quality content generation must be restricted to short, non-verbal, and quality-consistent data [49]. As existing ASMR datasets did not fulfill these requirements, we collected data by processing suitable ASMR YouTube videos. For a proof-of-concept, we focused on *tapping* sounds, which were found to be a well-functioning ASMR trigger for many people [2]. We used tapping-only videos of content creators well-known to the ASMR community, namely "ASMR Bakery"[2], "LottieLoves ASMR"[3], "Coromo Sara"[4] and "ricarda"[5]. We downsampled the audio to 16kHz mono using FFMPEG to ease the GAN's learning task. One limitation of GANs is that they require a fixed output length definition. Therefore, we extracted 2,000 audio segments of 1 second length using a threshold-based segmentation. Further, we applied *Fade-In* and *Fade-Out* effects to the segments. By doing so, we got rid of unwanted cracking or popping sounds at the start or end of the audio sample that the generator might otherwise reproduce. We divided the dataset into training, testing, and validation sets (70% train, 20% val, 10% test) to ensure robust evaluation and validation of the GAN model.

[2]https://www.youtube.com/watch?v=sIgkTYTWPz8
[3]https://www.youtube.com/watch?v=kMvGsOrpjNo
[4]https://www.youtube.com/watch?v=y03M_isyV3E
[5]https://www.youtube.com/watch?v=EGgKJsuM7Ns

3.1.2 GAN Training. We used the WaveGAN architecture presented by Donahue et al. [8]. WaveGAN is a modification of the DCGAN [34] architecture, which in turn extends the original GAN framework by replacing the fully connected layers by convolutional and deconvolutional layers. WaveGAN modifies the DCGAN to work well specifically with audio data, e.g., by replacing 2D kernels with 1D kernels and other adapatations. For a full architecture description, please refer to the WaveGAN publication [8]. The model was trained from scratch using the exact training configuration and hyperparameters recommended by Donahue et al. [8]. After 20k training steps, the generated audio samples resembled tapping sounds with various surfaces like glass and wood, and no further improvement could be observed. Minor artifacts were still present, but the overall results were satisfactory.

3.2 User Interface

Schlagowski et al. [37] compared different graphical user interfaces, called *Vector Manipulation Modules* (VMMs), that make audio generated by WaveGAN customizable. Opting for a VMM design with high hedonistic UEQ ratings in *Attractiveness*, *Stimulation*, and *Novelty*, we adopted one of their designs representing elements of the latent input space in 3D spheres (see Fig. 1). The user can drag and drop these spheres to adjust their corresponding numerical value in the latent input space of the GAN and hear the resulting change of the generated audio sample in real-time. As the WaveGAN model that we trained transforms a 100-dimensional noise input vector to audio data resembling ASMR tapping sounds, there are 100 manipulatable spheres in the GUI. The numerical latent space value is proportional to the spatial distance of the dragged (smaller) sphere to an inner and outer sphere, representing the minimum and maximum thresholds of the corresponding latent space value (semitransparent spheres in Fig. 1). The initial 3D positions are calculated using the fibonacci lattice algorithm [43].

Other application features include controls to add a *Reverb* effect, control the volume, and completely reshuffle the current input vector. Further, an additional *Size* slider is part of the system. By moving that slider, the number of controls for the input vector can be decreased so that the user can focus on only a few dimensions instead of all 100.

We embedded both the VMM and a Javascript-based version of WaveGAN into a webpage that uses the browser to display the VMM, run the GAN, and playback generated audio. Additional features of the resulting demonstrator system include:

(1) Playback of the last generated ASMR sound in a loop.
(2) An optional "Hands off" mode, in which the input vector automatically changes each time the sound is played within the loop.
(3) A tutorial mode, which explains the system by pointing to certain AI elements and describing their features.

4 EVALUATION

To evaluate if our application is suitable to induce an ASMR experience, we conducted an explorative user study in which participants had to interact with the ASMRcade system. Therefore, we made ASMRcade available on a web server and invited online participants to test it.

4.1 Study Group

We wanted to ensure that a good portion of our participants were already familiar with and actively engaging with the concept of ASMR. Thus, we targeted the existing ASMR community by creating a post in the subreddit "r/ASMR", an online platform where people share ASMR-related experiences and information on a regular basis. There, we briefly explained the application and invited users to test the ASMRcade application and complete the associated questionnaires. As a result, we acquired 17 Participants, which primarily consisted of users from the "r/ASMR" subreddit and others who discovered the Reddit post.

4.2 Study Procedure

During the study, the participants could first interact with the ASMRcade system for as long as they liked. After that, they had to complete an online questionnaire which consisted of three different parts:

(1) **Personal information:** This part contained three questions regarding age, gender, and familiarity with ASMR to understand the demographic profile of the participants.
(2) **User experience questionnaire:** We used the user experience questionnaire (UEQ) [23] to assess how participants experienced the interaction with the system. The questionnaire contains 26 items measuring six dimensions of user experience: *Attractiveness*, *Perspicuity*, *Efficiency*, *Dependability*, *Stimulation*, and *Novelty*. Each item is represented by two opposing terms and a corresponding seven-point scale ranging from -3 (most negative) to 3 (most positive), with 0 indicating a neutral response. To maintain the integrity of the online questionnaire, we followed the UEQ Handbook's advice to remove suspicious data. As such, participant responses were considered suspicious if:
   - They gave vastly different answers for questions within the same scale, and this occurred in at least three scales.
   - They provided identical answers to more than 15 of the 26 questions.
(3) **ASMR-specific questions:** The final section consisted of questions related to the participants' experience of ASMR while using ASMRcade. These questions aimed to gather information about whether participants experience ASMR in general, whether they experienced it while using ASMRcade, factors influencing the triggering or lack thereof, the time spent with the ASMRcade application, suggestions for improvement, likes and dislikes, opinions on the generated sounds and the interaction with the system. Specifically, we asked the following questions:
   - Q1: *Do you experience ASMR in general?* (Binary choice)
   - Q2: *Have you experienced ASMR while using the ASMRcade?* (Binary choice). Depending on their answer, participants had to provide free-form feedback on how they interacted with the application when ASMR was triggered (Q2.1) or what they thought would have to change to experience ASMR (Q2.2).
   - Q3: *How much time (in minutes) have you spent in the ASMRcade?* (Numerical-only free-form input)

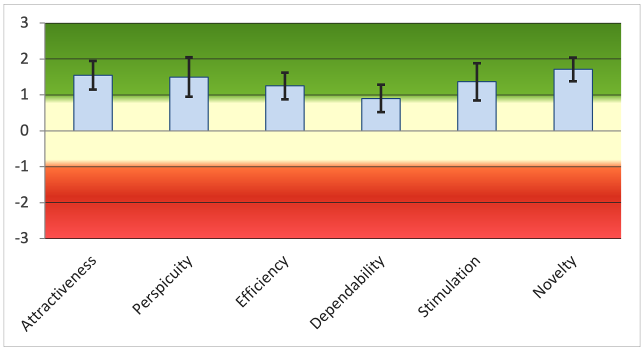

Figure 2: Mean values and variances for UEQ scales. The plot was created with the UEQ Data Analysis Tool, Version 12.

- Q4: *What improvements would you suggest for a better ASMR experience?* (Free form input)
- Q5: *What did you like about the ASMRcade?* (Free form input)
- Q6: *What did you dislike about ASMRcade?* (Free form input)
- Q7: *How did you like the generated sounds?* (Free form input)
- Q8: *How did you like the interaction with the system?* (Free form input)

4.3 Results

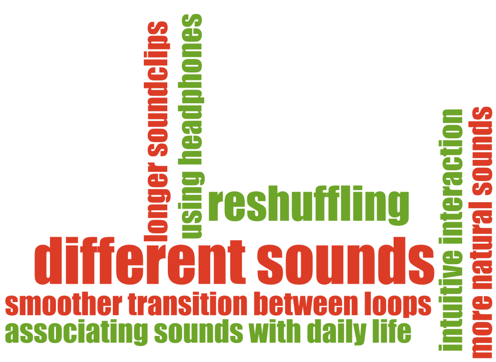

Figure 3: Word cloud resembling the responses to how the users were interacting with the application when ASMR was triggered (Q2.1, green), and what they think would need to change in order to experience ASMR (Q2.2, red).

In the following sections, we present the results of the online study. A total of 17 participants fully completed the survey. However, two participants reported spending 0 minutes with ASMRcade, leading to their removal before data analysis.

4.3.1 Demographic Questions. The demographic analysis of the remaining 15 participants revealed the following:

- Age: Participants ranged from 19 to 35 years old, with a mean age of 24.53 years.

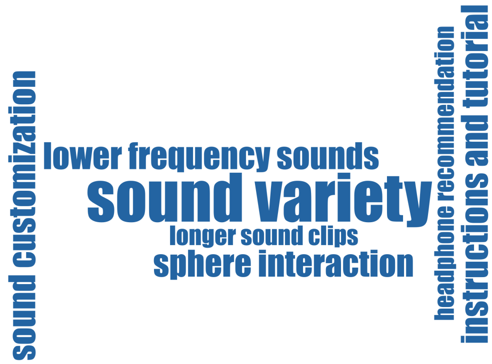

Figure 4: Improvement suggestions from participants (Q4).

- Gender: The sample included 10 males, 4 females, and 1 participant who preferred not to answer.
- Familiarity with ASMR:
  - 1 participant never heard about it before
  - 2 participants heard about it but never consumed ASMR content
  - 11 participant consumed ASMR content before
  - 1 participant consumed ASMR content occasionally
  - 0 participants consumed ASMR content regularly

4.3.2 User Experience Questionnaire (UEQ). To maintain the integrity of the questionnaire, we removed suspicious data as suggested in the UEQ Handbook's guidelines. As such, two data entries were removed, leaving 13 valid samples for analysis. We used the analysis tool provided by Laugwitz et al. [23] to analyze participants' responses. The results for the UEQ scales are presented in Table 1 and depicted in Figure 2.

The best-performing scale was Novelty (mean score of 1.712), followed by Attractiveness (1.551) and Perspicuity (1.5). Efficiency and Stimulation scored 1.25 and 1.365, respectively. The lowest-performing measure was Dependability, with a mean score of 0.904. However, according to the UEQ handbook, mean scores above 0.8 represent a *good evaluation*. Our application surpasses that value across all dimensions.

Scale	Mean	Variance
Attractiveness	1.551	0.54
Perspicuity	1.500	1.02
Efficiency	1.250	0.47
Dependability	0.904	0.50
Stimulation	1.365	0.90
Novelty	1.712	0.36

Table 1: UEQ Scales (Mean and Variance)

4.3.3 ASMR related questions. In the following, we present the results of the ASMR-specific questions from the online questionnaire. These questions aimed to assess participants' general ASMR experiences and their experiences with ASMRcade.

Figure 5: Word cloud resembling the responses to what users liked (Q5).

- General ASMR Experience: 8 participants reported that they generally experience ASMR.
- ASMR Experience in ASMRcade: 7 participants indicated that they had experienced ASMR while interacting with ASMRcade.
- ASMR Experience Overlap: Among the 8 participants who reported general ASMR experiences, five also experienced ASMR while using ASMRcade. ASMRcade triggered an ASMR experience for two users that generally don't experience ASMR.
- Time Spent in ASMRcade: On average, participants spent 12.47 minutes interacting with ASMRcade. The minimum reported time was 3 minutes, and the maximum was 30 minutes.

To analyze the free-form feedback from participants, an inductive thematic analysis [4] was conducted, which helps identify and report patterns within textual data. This qualitative method allows for understanding participants' experiences and opinions without imposing predetermined categories or theoretical perspectives and was previously used in similar studies (e.g., [7, 21, 38, 50]).

Code/word clouds were created using the themes and codes identified in the analysis to represent user feedback visually. Code clouds display words or phrases in varying font sizes, with larger sizes representing higher occurrence frequency within the data.

Figure 3 presents the code cloud to Q2.1 and Q2.2. Participants were shown one of those questions depending on how they answered regarding experiencing ASMR while interacting with ASMRcade. Green words represent feedback on how participants (that reported experiencing ASMR while interacting with ASMRcade) interacted with ASMRcade when their ASMR was triggered. Red words represent participants' feedback that they hadn't experienced ASMR while interacting with ASMRcade about what they thought was missing or needed to change for them to experience ASMR.

The feedback by users that had experienced ASMR (green) includes aspects such as intuitive interaction (1 mention; e.g., "Changing the sounds became way more intuitive"), reshuffling (2 mentions; e.g., "Reshuffling until I experienced it"), associating sounds with daily life (1 mention; e.g., "My brain started to find references from my daily life from which the sound could be known and tried to find

Figure 6: Word clouds resembling the responses to what users disliked (Q6)

Figure 7: Feedback to the generated sounds (Q7).

Figure 8: Feedback about the interaction experience (Q8).

rhythm and movement that could be applied to the sound effects"), and using headphones (1 mention).

The feedback on what was missing to experience ASMR (red) encompasses comments on the need for longer sound clips (1 mention), different sounds (3 mentions; e.g., "Other sounds" and "Different sounds to choose from, not just tapping"), more natural sounds (1 mention; e.g., "I would like the sounds to be more natural, they felt kind of mechanical in the ASMRcade"), and smoother transition between loops (1 mention; e.g., "The transition between the loops was pretty noticeable which made it feel less natural to me").

Figure 4 displays the code cloud visualizing participants' suggestions for improvements to enhance the ASMR experience (i.e., Q4). The codes encompass various improvement suggestions, such as sound variety (5 mentions; e.g., "Different sounds," "Bigger variety of sounds," and "Diversify sound categories"), instructions and tutorial (2 mentions; e.g., "Clear instructions" and "Clickthrough tutorial"), sound customization (2 mentions; e.g., "Changing the speed"), longer sound clips (1 mention), lower frequency sounds (2 mentions; e.g., "Sound frequency adjustment" and "More profound bass sounds"), sphere interaction (2 mentions; e.g., "Sphere influence visualization" and "3D model interface improvement"), and headphone recommendation (1 mention).

Figure 5 presents the code cloud illustrating the aspects participants liked about ASMRcade (i.e., Q5). The codes highlight various positive aspects, including sound variety (4 mentions; e.g., "Endless sound possibilities," "Different patterns," and "Wide variety of sounds"), interface and usability (6 mentions; e.g., "Intuitive interaction," "Nice interface," "Minimalistic interface," and "Easy to use"), sound customization (5 mentions; e.g., "Volume and reverb control," "Realistic sounds," "Customizable sounds," and "User control over sounds"), reshuffle button (1 mention), browser compatibility (1 mention; e.g., "Runs in-browser"), vector representation (1 mention; e.g., "Vector representation was quite nice"), engaging experience (4 mentions; e.g., "Engaging and creative," "Interactive experience," and "Unique experience"), and visual design (1 mention; e.g., "Visually pleasing").

Figure 6 presents the code cloud for user feedback regarding the aspects they disliked about ASMRcade (i.e., Q6). Unintuitive tutorial button (1 mention; e.g., "I disliked the question mark button as a Tutorial being very small."), Understanding size change (2 mentions; e.g., "I didn't understand the size change. I couldn't hear any difference with changing the size"), Unknown sphere function (3 mentions; e.g., "I didn't really understand what manipulating the spheres actually does", "fine-tuning is a little bit confusing"), Small lags (1 mention), Overlapping text (2 mentions; e.g., "bubble thingy laid over the text"), length of sound clips (1 mention), high frequencies (1 mention), Repetitive sounds (2 mentions; e.g., "Some sounds seemed to appear repetitively even though different orbs were manipulated", Limited sound variety (3 mentions; e.g., "too few different sounds", "it can only generate relatively unexciting sounds (i.e. tapping)"), Visually unpleasing (1 mention).

Figure 7 presents the code cloud generated based on participants' feedback on the generated sounds in ASMRcade (i.e., Q7). The positive feedback (green) includes aspects such as sound quality (4 mentions; e.g., "The sounds were pleasant" and "The tapping sounded convincing"), sound realism (2 mentions; e.g., "The knocking sounds were pleasant"), good sound variety (1 mention), reverb effect (1 mention) and unique experience (1 mention; e.g., "The loops added a unique and enjoyable aspect to the experience").

On the other hand, negative feedback (red) encompasses comments on the need for greater sound variety (3 mentions; e.g., "Some other sounds would've been nice" and "More variety"), loop transitions (2 mentions; e.g., "Regulate the length of the sounds" and "The transition between loops felt disruptive"), and unnatural sounds (1 mention; e.g., "They didn't feel very natural to me").

Figure 8 presents the code cloud for user feedback regarding their interaction with ASMRcade (i.e., Q8). It is composed of three colors: green for positive feedback, red for negative feedback, and yellow for neutral feedback. The codes highlight different aspects, including Enjoyable experience (3 mentions; e.g., "It was an enjoyable experience", "It was interesting and was fun to use"), Intuitive interaction (5 mentions; e.g., "The interaction was easy and intuitive", "It was mostly intuitive", "Easy to understand, learning by doing worked fast"), Tutorial issue (2 mentions; e.g., "the Tutorial problem", "Would have liked to be able to pause/click through the tutorial myself"), Unique interaction features (2 mentions; e.g., "I was pleasantly surprised that the system placed the balls at the maximum possible distance when I moved them too far", "Nice imaging of the different bubbles as an input vector for the neuronal network"), Interface issues (1 mention; e.g., "Bubbles partly laid over the text"), Potential for improvement (2 mentions; e.g., "More possibilities to customize and play around with different sounds would make it very enjoyable").

5 DISCUSSION

User Experience. Overall, the User Experience Questionnaire (UEQ) results indicate a positive user experience with ASMRcade. The highest mean scores were observed for Novelty, Attractiveness, and Perspicuity, suggesting that ASMRcade was perceived as an innovative and appealing solution with a clear and understandable design. In the open feedback, one user stated, "It's a very interesting experience that I haven't had like that before. I really enjoyed customizing the sounds to what I wanted to hear, which is something I can't do when I'm watching an ASMR video." The relatively lower scores for Efficiency, Stimulation, and Dependability indicate room for improvement in these more pragmatic and usability-related scores. The efficiency and dependability of the application might be affected by the inherent unpredictability of WaveGAN. Users may need to manipulate multiple spheres to produce a pleasant sound, which might need to be more efficient. One user mentioned, "I felt like manipulating the spheres didn't really change much".

Overall, as the Pragmatic Quality scales (Perspicuity, Efficiency, Dependability) had lower mean scores than the Hedonic Quality scales (Stimulation, Originality), one can conclude that participants found ASMRcade to be more enjoyable and stimulating than efficient and dependable. As such, ASMRcade might not be reliable for generating desired sounds on every interaction because of the GANs' inherent unpredictability. This finding suggests that while the application offers an engaging and novel experience, its practical aspects may need further refinement to enhance user satisfaction. However, we also note that goal-oriented qualities may be less critical for an application designed for relaxation and de-stressing, such as ASMRcade. Furthermore, higher scores in Hedonic qualities may be attributed to the use case itself, as experiencing ASMR is a relaxing and immersive experience in itself [1, 33]. One participant's open feedback supports this notion: "It created curiosity and started to engage the creative mind."

ASMR Experiences. More than half of the participants reported generally experiencing ASMR, with almost the same proportion experiencing ASMR while interacting with ASMRcade. The overlap between these groups demonstrates ASMRcade's potential to induce ASMR in users susceptible to the phenomenon. However, not all participants who generally experienced ASMR reported experiencing it with ASMRcade, suggesting that the application may need further improvements to cater to a broader range of ASMR triggers and preferences. On the other hand, certain participants were able to experience ASMR that did not experience ASMR before. This observation indicates that the interactive component we introduced through the ASMRcade application might have advantages over passive ASMR consumption. Interacting with an ASMR content generator can even be seen as a new category of ASMR triggers, which, like other categories, may work very well for some users but less for others. Here, future work has to dive more into detail on how exactly interactive and passive ASMR consumption differs in triggering ASMR experiences. The open feedback collected from participants provided valuable insights into their experiences and opinions regarding ASMRcade, which complemented the quantitative data obtained from the User Experience Questionnaire (UEQ). By applying an inductive thematic analysis approach to the open feedback, several recurring themes were identified that shed light on the strengths and areas for improvement of the application. One of the key strengths of ASMRcade highlighted by the participants was the intuitive interaction. Participants found navigating and manipulating the spheres to create different sounds easy. For example, one participant noted that "Changing the sounds became way more intuitive" once they got the hang of the system. This feedback corresponds with the above-average scores for Attractiveness and Perspicuity in the UEQ results. However, participants also mentioned that some aspects of ASMRcade's sphere functions needed clarification. One user stated, "I didn't really understand what manipulating the spheres actually does," while another mentioned, "Fine-tuning is a little bit confusing." These comments suggest that some users struggled to comprehend how the sphere manipulation impacted the generated sounds. Addressing this issue by providing clearer explanations about the function and limits of sphere interaction, visual cues, or tooltips could help users better understand the connection between sphere manipulation and the resulting sounds, thus enhancing their overall experience. One way to give visual cues would be to do a Vector Impact Analysis, like Schlagowski et al. did for their drum sequencer [37]. After analyzing their impact on the generated sounds, spheres can be sorted accordingly, and shaders could be used to indicate individual spheres' impact on different frequency bands. The open feedback also revealed that participants appreciated the sound customization options available in ASMRcade but desired more advanced customization features. Users found the ability to manipulate volume and reverb controls helpful, as one participant noted, "I liked the option that you could separately change the volume and the Reverb" However, some users expressed a desire for even more control over sound characteristics, such as speed and frequency adjustments. For example, one participant suggested, "I would also like there to be a slider where I could manipulate the speed of the sound". Further constructive feedback on areas that could be improved was provided. For instance, several users mentioned that the application would benefit from a more diverse selection of sounds. One participant suggested, "Diversify the sound categories". Another user mentioned that the sound transitions could be smoother, stating, "The transition between the loops was pretty noticeable, which made it feel less natural to me." Furthermore, some participants found certain features to need to be clarified or to be more intuitive within ASMRcade. For example, one user noted that the tutorial button needed to be bigger and easier to notice. Another participant mentioned difficulties understanding the impact of changing the sphere size slider, which hides certain manipulatable spheres: "I didn't understand the size change; I couldn't hear any difference with changing the size." To address these issues, the application could be refined by improving the visibility and accessibility of the tutorial, as well as providing more precise explanations of what the size slider does and how it affects the generated sounds. As the by far most recurring theme was the diversity of sounds, the logical next step should be to train the WaveGAN not only on tapping sounds but on a broad spectrum of different ASMR audio triggers that exist. These could include scratching, whispering, or other mouth-made sounds. By addressing these areas, ASMRcade could further enhance user satisfaction and provide a more engaging and enjoyable experience for users seeking relaxation and stress relief through ASMR. Further, it should be noted that our exploratory user study had a relatively small sample size. Future research has to delve more into detail by studying specific characteristics of interactive ASMR that contribute to the users' experience through quantitative user studies and more participants.

6 CONCLUSION

In this work, we presented ASMRcade, an application for interactively exploring and generating personalized ASMR audio triggers. Using this application, users could change the input vector for a WaveGAN capable of transforming that input vector to ASMR-triggering tapping sounds by manipulating the synthesized sounds via a web-based user interface. A first explorative user study indicates that some users can benefit from the interactive approach and that it might be able to broaden accessibility for ASMR experiences as some users had their first successful ASMR triggers with our system.

From user experience questionnaires and the qualitative feedback, we conclude that users appreciated the intuitive interaction with the application, which was also perceived as innovative and appealing. Users especially enjoyed the sound customization options, but requested further control over the speed and frequency of the generated sound. Furthermore, while the sound quality was good enough, it could be further improved by using a higher sampling rate and stereo signals. Also, many users asked for more variety in the generated sounds. As such, future work has to extend the system to a larger set of ASMR sound categories.

To conclude, the study is a proof of concept for a novel interactive approach for generating ASMR triggers. The substantial share of users that experienced ASMR and their interest in improving ASMRcade shows great potential for an interactive ASMR experience as a suitable alternative to conventional ASMR content.

ACKNOWLEDGMENTS

This paper was partially funded by the Bavarian Ministry of Science and Arts (Bavarian Research Association ForDigitHealth).

REFERENCES

[1] Emma L. Barratt and Nick J. Davis. 2015. Autonomous Sensory Meridian Response (ASMR): a flow-like mental state. *PeerJ* 3 (2015), e851. https://doi.org/10.7717/peerj.851

[2] Emma L. Barratt, Charles Spence, and Nick J. Davis. 2017. Sensory determinants of the autonomous sensory meridian response (ASMR): understanding the triggers. *PeerJ* 5 (2017), e3846. https://doi.org/10.7717/peerj.3846

[3] Axel Berndt, Nadia Al-Kassab, and Raimund Dachselt. 2015. TouchNoise: A New Multitouch Interface for Creative Work with Noise. In *Proceedings of the Audio Mostly 2015 on Interaction With Sound*. 1–8.

[4] Virginia Braun and Victoria Clarke. 2012. *Thematic analysis.* 57–71.

[5] Mihaly Csikszentmihalyi. 2000. *Beyond boredom and anxiety.* Jossey-Bass.

[6] Palle Dahlstedt. 2001. Creating and exploring huge parameter spaces: Interactive evolution as a tool for sound generation. In *ICMC*. Citeseer.

[7] Daniel Diethei, Jasmin Niess, Carolin Stellmacher, Evropi Stefanidi, and Johannes Schöning. 2021. Sharing heartbeats: motivations of citizen scientists in times of crises. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–15.

[8] Chris Donahue, Julian McAuley, and Miller Puckette. [n. d.]. Adversarial Audio Synthesis. http://arxiv.org/pdf/1802.04208v3

[9] Chris Donahue, Julian McAuley, and Miller Puckette. 2018. Adversarial audio synthesis. *arXiv preprint arXiv:1802.04208* (2018).

[10] Hao-Wen Dong, Wen-Yi Hsiao, Li-Chia Yang, and Yi-Hsuan Yang. 2018. Musegan: Multi-track sequential generative adversarial networks for symbolic music generation and accompaniment. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32.

[11] Jesse Engel, Kumar Krishna Agrawal, Shuo Chen, Ishaan Gulrajani, Chris Donahue, and Adam Roberts. 2019. Gansynth: Adversarial neural audio synthesis. *arXiv preprint arXiv:1902.08710* (2019).

[12] Zexin Fang, Bin Han, C. Clark Cao, and Hans. D. Schotten. [n. d.]. Artificial ASMR: A Cyber-Psychological Study. https://doi.org/10.48550/arxiv.2210.14321

[13] Daniela Gasperin, Gopalakrishnan Netuveli, Juvenal Soares Dias-da Costa, and Marcos Pascoal Pattussi. 2009. Effect of psychological stress on blood pressure increase: a meta-analysis of cohort studies. *Cadernos de saude publica* 25, 4 (2009), 715–726.

[14] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial networks. *arXiv preprint arXiv:1406.2661* (2014).

[15] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative Adversarial Networks. https://doi.org/10.48550/ARXIV.1406.2661

[16] George Halkos and Dimitrios Bousinakis. 2010. The effect of stress and satisfaction on productivity. *International Journal of Productivity and Performance Management* 59, 5 (2010), 415–431.

[17] Vladimir Iashin and Esa Rahtu. 2021. Taming visually guided sound generation. *arXiv preprint arXiv:2110.08791* (2021).

[18] Maximos Kaliakatsos-Papakostas, Aggelos Gkiokas, and Vassilis Katsouros. 2018. Interactive control of explicit musical features in generative lstm-based systems. In *Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion*. 1–7.

[19] Ronald C Kessler. 1997. The effects of stressful life events on depression. *Annual review of psychology* 48, 1 (1997), 191–214.

[20] Josephine Klefeker, Libi Striegl, and Laura Devendorf. 2020. What HCI can learn from ASMR: Becoming enchanted with the mundane. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[21] Veronika Krauß, Alexander Boden, Leif Oppermann, and René Reiners. 2021. Current practices, challenges, and design implications for collaborative AR/VR application development. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–15.

[22] Yuyo Lai, Shyh-Kang Jeng, Der-Tzung Liu, and Yo-Chung Liu. 2006. Automated optimization of parameters for FM sound synthesis with genetic algorithms. In *International Workshop on Computer Music and Audio Technology*. Citeseer, 205.

[23] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and Evaluation of a User Experience Questionnaire. In *HCI and Usability for Education and Work*, Andreas Holzinger (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 63–76.

[24] Chae Young Lee, Anoop Toffy, Gue Jun Jung, and Woo-Jin Han. 2018. Conditional wavegan. *arXiv preprint arXiv:1809.10636* (2018).

[25] Bruce S McEwen. 2008. Central effects of stress hormones in health and disease: Understanding the protective and damaging effects of stress and stress mediators. *European journal of pharmacology* 583, 2-3 (2008), 174–185.

[26] Silvan Mertes, Alice Baird, Dominik Schiller, Björn W Schuller, and Elisabeth André. 2020. An Evolutionary-based Generative Approach for Audio Data Augmentation. In *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*. IEEE, 1–6.

[27] Eduardo Reck Miranda and John Al Biles. 2007. *Evolutionary computer music.* Springer.

[28] Thomas Mitchell. 2012. Automated evolutionary synthesis matching: Advanced evolutionary algorithms for difficult sound matching problems. *Soft Computing* 16 (2012), 2057–2070.

[29] Dan Morris, Ian Simon, and Sumit Basu. 2008. Exposing parameters of a trained dynamic model for interactive music creation. (2008).

[30] Zhen Nan and Makoto Fukumoto. 2022. ASMR Sound Generation Simulating the Sounds Heard by a Fetus Using Interactive Evolutionary Computation. In *2022 Joint 12th International Conference on Soft Computing and Intelligent Systems and 23rd International Symposium on Advanced Intelligent Systems (SCIS&ISIS)*. IEEE, 1–4.

[31] Ji Yeon Oh, Daun Kim, Jae-Yeop Jeong, Jin-Woo Jeong, and Elizaveta Lukianova. 2023. Tingle Just for You: A Preliminary Study of AI-based Customized ASMR Content Generation. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–7.

[32] Markku Partinen. 1994. Sleep disorders and stress. *Journal of Psychosomatic Research* (1994).

[33] Giulia Lara Poerio, Emma Blakey, Thomas J. Hostler, and Theresa Veltri. 2018. More than a feeling: Autonomous sensory meridian response (ASMR) is characterized by reliable changes in affect and physiology. *PloS one* 13, 6 (2018), e0196645. https://doi.org/10.1371/journal.pone.0196645

[34] Alec Radford, Luke Metz, and Soumith Chintala. [n. d.]. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. http://arxiv.org/pdf/1511.06434v2

[35] António Ramires, Jordan Juras, Julian D Parker, and Xavier Serra. 2022. A STUDY OF CONTROL METHODS FOR PERCUSSIVE SOUND SYNTHESIS BASED ON GANS. (2022).

[36] Hannes Ritschel, Ilhan Aslan, Silvan Mertes, Andreas Seiderer, and Elisabeth André. 2019. Personalized synthesis of intentional and emotional non-verbal sounds for social robots. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 1–7.

[37] Ruben Schlagowski, Silvan Mertes, and Elisabeth André. 2021. Taming the Chaos: Exploring Graphical Input Vector Manipulation User Interfaces for GANs in a Musical Context. In *Audio Mostly 2021*. ACM, New York, NY, USA, 216–223. https://doi.org/10.1145/3478384.3478411

[38] Ruben Schlagowski, Dariia Nazarenko, Yekta Can, Kunal Gupta, Silvan Mertes, Mark Billinghurst, and Elisabeth André. 2023. Wish You Were Here: Mental and Physiological Effects of Remote Music Collaboration in Mixed Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*.

[39] Ruben Schlagowski, Fabian Wildgrube, Silvan Mertes, Ceenu George, and Elisabeth André. 2022. Flow with the Beat! Human-Centered Design of Virtual Environments for Musical Creativity Support in VR. In *Creativity and Cognition*. ACM, New York, NY, USA, 428–442. https://doi.org/10.1145/3527927.3532799

[40] Diemo Schwarz. 2005. Current research in concatenative sound synthesis. In *International Computer Music Conference (ICMC)*. 1–1.

[41] Yujun Shen and Bolei Zhou. 2021. Closed-form factorization of latent semantics in gans. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 1532–1540.

[42] Jeff Snyder and Danny Ryan. 2014. The Birl: An Electronic Wind Instrument Based on an Artificial Neural Network Parameter Mapping Structure.. In *NIME*. 585–588.

[43] Richard P Stanley. 1975. The Fibonacci lattice. *Fibonacci Quart* 13, 3 (1975), 215–232.

[44] Ric G Steele, Jeffrey A Hall, and Jennifer L Christofferson. 2020. Conceptualizing digital stress in adolescents and young adults: Toward the development of an empirically based model. *Clinical Child and Family Psychology Review* 23 (2020), 15–26.

[45] Matthew A Stults-Kolehmainen and Rajita Sinha. 2014. The effects of stress on physical activity and exercise. *Sports medicine* 44 (2014), 81–121.

[46] Jérôme Sueur, Thierry Aubin, and Caroline Simonis. 2008. Seewave, a free modular tool for sound analysis and synthesis. *Bioacoustics* 18, 2 (2008), 213–226.

[47] Robert Tubb and Simon Dixon. 2014. The Divergent Interface: Supporting Creative Exploration of Parameter Spaces.. In *NIME*. 227–232.

[48] Emily C Weinstein and Robert L Selman. 2016. Digital stress: Adolescents' personal accounts. *new media & society* 18, 3 (2016), 391–409.

[49] Karren Yang, Bryan Russell, and Justin Salamon. 2020. Telling left from right: Learning spatial correspondence of sight and sound. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 9932–9941.

[50] Tianyi Zhang, Björn Hartmann, Miryung Kim, and Elena L Glassman. 2020. Enabling data-driven api design with community usage data: A need-finding study. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.