Article

Can Voice-Responsive VR Art Experiences Improve Wellbeing?

ADINDA, R. VAN 'T KLOOSTER, Manchester Metropolitan University, United Kingdom;
email: A.vant.Klooster@mmu.ac.uk

ROBYN, E. L. DOWLEN, Edge Hill University, United Kingdom;
email: robyn.dowlen@edgehill.ac.uk

10.2478/bsmr-2025-0010

ABSTRACT

This paper outlines the design journey of two VR artworks, *AudioVirtualizer* (2019) and *VRoar* (2023), to inform the development of future VR artwork. A central premise of the first two interfaces is that abstract graphics are manipulated by live sound in VR: the *AudioVirtualizer* uses pre-composed music and soundscapes or microphone input to move and deform graphics (van 't Klooster & Collins, 2021) and *VRoar* is fully voice-responsive (van 't Klooster, & Collins, 2024). From informally observing people interact with these interfaces, we suspected through participant reflections that there had been an increase in positive emotions obtained from the VR art experience. To validate this, we evaluated both interfaces using a questionnaire before and after the experience, during an exhibition in the Holden Gallery in Manchester and across various venues in the Northeast of England. Forty-one participants experienced VRoar and fifty people experienced the AudioVirtualizer. From the evaluation we learned that a short-term increase in positive emotions is achievable from designed VR art experiences. The *AudioVirtualizer* achieved this for the group of non-trained singers and the *VRoar* interface achieved it for trained singers. Singers and non-singers engaged differently, suggesting a tailored design is required. This paper presents qualitative and quantitative findings to highlight the impact on participants' wellbeing as well as how feedback will inform the design of the next interface.

KEYWORDS: Emotion Sensitive Art, Virtual Reality, wellbeing, responsive landscapes, audio machine learning

INTRODUCTION

Virtual Reality (VR) offers a rapidly evolving medium for artistic experience and expression, yet its potential for voice- and body movement-based interaction remains underexplored to date. Using the voice as a controller not only opens creative possibilities but also enhances accessibility in engaging with software and interactive interfaces – particularly for individuals who may have limited mobility. Beyond its functional role, the voice is closely tied to personal identity (Edgar 2019) and emotional wellbeing (Kang et al, 2018). While some might feel initially self-conscious about vocal expression (Harma, 2011), engaging the voice in creative contexts can offer opportunities for therapeutic gain (Clift, 2012; Berluche, 2024), which can enhance an individual's sense of wellbeing. While there is a growing interest in the use of VR experiences to enhance mental health and wellbeing (Wagener et al, 2021), there has been little exploration of the specific role of the voice in complementing the experience. In *Mood Worlds,* Wagener et al (2022) use artistic expression as a way to increase positive emotions in the participant, arguing that the user's sense of agency is boosted by being allowed to create their own emotional world using brushes, which in turn improves their wellbeing. In our VR art experience *VRoar* (2023)*,* we do something similar, only in this case the available means of expression is the voice which in turn generates and affects graphics. The past year has seen an explosion of VR apps that use vocal input and AI to have virtual humans respond in virtual spaces, and many of these apps claim to improve wellbeing. However, this claim is often made without robust evaluation, leaving the therapeutic benefit questionable.

In VR art, the use of sound responsiveness has been only infrequently explored. In 2018, Marshmallow Laser Feast created a mixed reality installation featuring an interactive part where users inside can generate 'voice fruits' using their voice. These are then dispersed when a bowling ball outside is thrown onto a massive gong (Gosling 2018). In *Re-Animated* (2022) by Jakob Kudsk Steensen, plants, moss, and insects respond to the pulse of the live generated music, and the audience's breath and voice influence the graphics through the microphone of the VR headset, though this is only a side detail in the work (Bezan, 2022). However, in *Deep Listener*, an Augmented Reality artwork from 2019 by the same artist this sound responsive element is used again (Steensen, 2025). (There are also biofeedback VR artworks, such as *Respire* by Prpa et al. (2018), where the user's breathing determines the positioning of the participant in a sea of water. This work is reminiscent of *Ephémère,* made in 1998 by VR art pioneer Char Davies (2005), which is much more complex visually but also steered by the immersant's breathing and other biosignals.

One relevant existing example of a live voice-responsive artwork that uses AI is *Miraj* by SOSO (2003-2025). Originally built as an app for AppleTV (SOSO, 2016), it was later installed in a DOME structure in a children's museum in San Diego. This system detects the meaning of individual nouns, and creates a live kaleidoscopic image based on the main detected noun. Whilst clearly attractive for children as a responsive sensory environment, this system is not in VR. Other than replacing the core sprite for each projection, the system does not display a lot of variety. An older but related example that is more rule-based and does not use AI is Golan Levin's *Messa di Voce* (2003), originally built for two vocalists but later adapted for gallery installation. It spurs theatrical abstract animated graphics to encourage the user to play with their voice. Whilst again not in VR but rather simply 2D, this installation is engaging and playful, especially when the graphics seem to visually emerge from the voice actor's mouths.

Ventriloquy I and II by Bryan Dunphy (2022) are marginally related to our work in that Dunphy likewise tries to create live audiovisuals that can be experienced in VR, but his work is different in that he alternates between mapping from sound to image or from image to sound and frequently presents the work as projections, rather than as VR art. There is the wider context of live generated audiovisuals, with many examples of audio generated graphics and, over the years, multiple different softwares and historically coloured organs, to support this, but these fall outside the remit of this paper as we are particularly interested in voice responsivity in a live VR context. There are previous VR audio-visualisers which generate graphics, such as plane9, with 250 different scenes the user can mix; FantaSynth; Chromestasia; Synestasia; Raybeam VR; CyberDreamVR, inspired by rave music (Weinel, 2019); and Vision, STEAM's own music visualiser in VR, released in 2018. Out of the abovementioned, only Chromestasia and Synestasia are responsive to live sound input such as offered in our *AudioVirtualizer* and *VRoar* interfaces, though Synestasia is more an app for VJ performers rather than a ready-made experience, has multiple pre-made scenes to choose between and has preset feature mappings that are not always entirely transparent. Kaleidosync is a third-party Spotify visualiser that syncs with a person's Spotify account and creates real-time, animated visualisations based on the music and parameters chosen by the user. VZX-Pro by Vovoid Media Technologies allows microphone input for the context of live audio performance with a live projection of animated graphics, but is not for VR.

In gaming, voice responsivity has been haltingly explored for decades. The first virtual pet game that used voice interaction was *Hey You Pikachu* in 1998 (Allison et al, 2020), and Karaoke has been a popular genre of voice interaction gaming, especially in the first decade of the 21st century (Allison et al, 2020). A recent example of voice control in a VR Game is the *Harry Potter VR Game,* which allows the user to cast spells by speaking certain spell words (BMFVR Gaming, 2022).

While voice responsiveness is still a novelty in the context of VR artworks, there has been a sharp rise in the number of HCI papers published on voice interaction (Deshmukh and Chalmeta, 2024). In the arts, scientific evaluation of the emotional effect of artworks, such as has been done here with *VRoar* and *AudioVirtualizer,* is also novel. This article combines artistic VR creation with quantitative and qualitative evaluation to gain more clarity on how voice-controlled VR artworks can improve wellbeing, and is the first systematic evaluation of voice-responsive VR art for affective outcomes. In this study, we view wellbeing not only as the presence of positive emotions, but as a broader, dynamic process that encompasses engagement, meaning, and self-expression (Butler & Kern, 2016). Within arts and health contexts, even short-term shifts in mood can represent meaningful moments of connection, release, or creative agency that contribute to an individual's overall sense of wellbeing. Evaluating these momentary changes offers valuable insight into how designed digital art experiences might foster emotional regulation and support mental wellbeing through creative engagement.

A useful way to capture the emotional impact of VR experiences is through the lens of positive and negative affect. These terms, widely used in psychology, describe broad dimensions of emotional experience with positive affect referring to the extent a person feels for example enthusiastic, active and alert, while negative affect reflects, amongst others, distress, fear, guilt or nervousness (Watson et al., 1988). Together, they represent relatively independent systems, meaning an individual can experience high levels of both (for example, excitement mixed with anxiety) or low levels of both (a neutral, flat mood). Tools such as the Positive and Negative Affect Schedule (PANAS, Watson et al., 1988) are frequently used to evaluate these dimensions and have been applied widely in the context of wellbeing studies. In the context of VR artworks, enhancing positive affect may be especially valuable since elevated positive emotions are associated with greater openness, creativity and resilience, all of which are linked to wellbeing (Friedrickson, 2001). By contrast, VR interactions could unintentionally heighten negative affect, for instance by making participants feel disorientated, frustrated, or self-conscious. Thus, considering both positive and negative affect within the context of an evaluation of VR artwork experiences offers a nuanced picture of how VR artworks may support wellbeing and offer up design choices that can be integrated into future interfaces that seek to enhance feelings of wellbeing. In this article, we explore this through the evaluation of two VR experiences: the *AudioVirtualizer* and *VRoar*, which are described in more detail below.

AudioVirtualizer (2019) is a VR installation on the HTC Vive Pro where abstract sculptures respond to music audio features, with multiple scenic environments that the user can access by finding and clicking on a four-legged sculpture (van 't Klooster & Collins, 2021). The *AudioVirtualizer* is the first author's initial investigation into VR, and the graphics are responsive to music which comes embedded within each of the scenes, but can also be changed via a menu option that gives four different soundscapes (two composed by Nick Collins and two by Adinda van 't Klooster) and a further option for microphone input. Each of the three different levels (after the entrance room, which is just for practising navigation) have different graphics and mapped behaviours which the user can observe and explore in relation to the different sound inputs (see Figures 1–3).

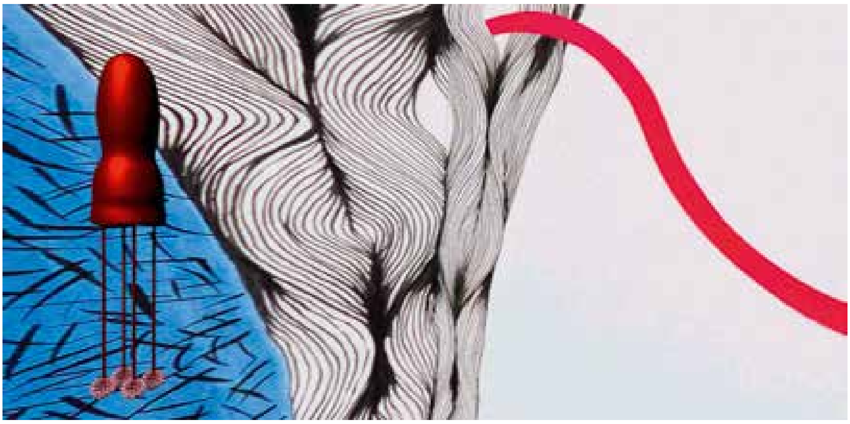
During a small-scale (unpublished) evaluation of the *AudioVirtualizer* in 2023 at NewBridge Studios in Newcastle-upon-Tyne, we observed an increased level of excitement in participants when they were able to visually manipulate graphics using their voice. This made us decide in the next work, *VRoar* (2023), to prioritise live vocal input rather than precomposed soundscapes to generate the visual distortions and animations (van 't Klooster and Collins, 2024). *VRoar* also runs on a wireless Quest headset rather than the wired HTC VIVE to allow for easier sharing and exhibition of the interface. What is shared across both interfaces is a similar hand-painted style for the backdrops, which are turned into skyboxes (see Figures 4 and 5); as well as the Unity plugin created by Nick Collins, which allows for the tracking of mid- and low-level sound features and facilitates mapping to control changing graphics.

What is also new in *VRoar* is that trees can be generated through navigating the landscape via movement of the joystick combined with the direction of one's gaze. Blobs appear with each detected onset and these blend to use a different texture when certain sound features change, whilst comic clouds distort and dance based on the sound the participant makes. Pressing the A-button swaps mappings again and brings in a new background image. All the elements in a scene pulse and distort through expressive voice control.

The previously mentioned evaluation at Newbridge also helped form the evaluation reported in this article, as we wanted to see if the informally noticed positive effect on mood could be scientifically substantiated. The present study therefore measures outcomes relating to the VR art experience as well as positive and negative emotions to understand whether there is any possibility for voice-responsive VR art to contribute to a person's sense of wellbeing. According to Seligman, positive emotion is one of the five pillars that determine wellbeing, alongside engagement, relationships, meaning and accomplishments (Butler and Kern, 2016). We focus specifically on the increase in positive emotion and the differences between singers (trained voice users) and non-singers, aiming to explore whether specific interface design adaptations may be required to enhance suitability and engagement for individuals with vocal or no vocal training.

METHODS

The two interfaces were tested by trained singers and non-trained vocalists. These groups were selected to enable a comparative analysis based on participants' familiarity with vocal techniques. The interfaces were exhibited in the Holden Gallery in Manchester and at various venues in the



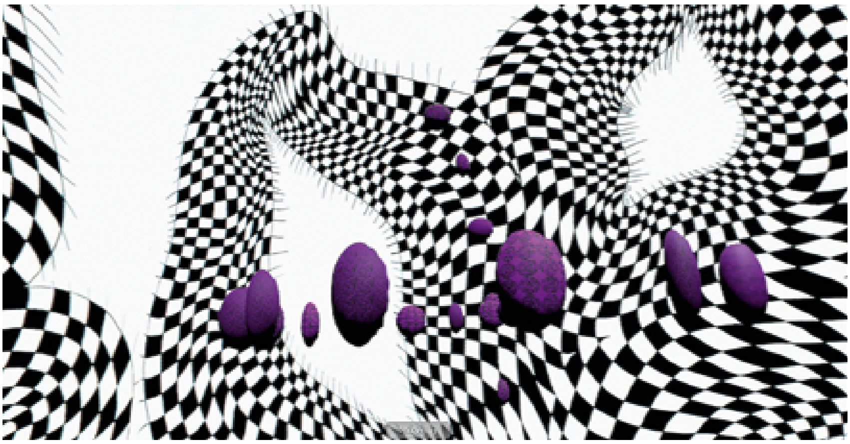


FIGURE 1. ***AudioVirtualizer*, entrance room showing the portal to the three levels, © Adinda van 't Klooster & Nick Collins, 2019**

FIGURE 2. ***AudioVirtualizer*, scene 1, © Adinda van 't Klooster & Nick Collins, 2019**

FIGURE 3. ***AudioVirtualizer*, scene 2, © Adinda van 't Klooster & Nick Collins, 2019**

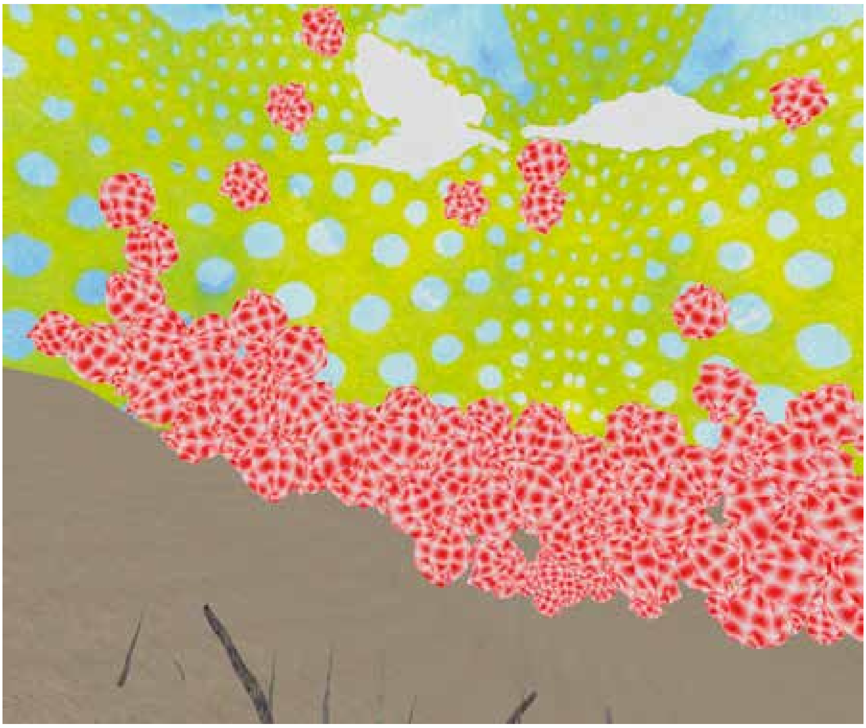


FIGURE 4. ***VRoar*, © Adinda van 't Klooster, 2023**

FIGURE 5. ***VRoar*, © Adinda van 't Klooster, 2023**

Northeast of England, with 41 participants experiencing *VRoar* and 50 experiencing *AudioVirtualizer*. A minority experienced both interfaces, but there were always at least three hours in between both experiences and the interface they experienced first was always random. We had 44 singers and 47 non-singers undertake our survey. Within the singer group, 64% were self-reported professionals or semi-professional singers and 34% were amateurs or beginners, while 2% were professionals but not currently practising (This breakdown is solely descriptive and not used for further inferential testing.) Participants were recruited through emailing staff and students at Manchester Metropolitan University and through contacting collaborator networks to engage with the interfaces onsite and take part in the associated evaluation.

Quantitative data – PANAS

The Positive and Negative Affect Scale (PANAS, Watson and Clark, 1988) was used to measure change in emotional wellbeing before and after the experience. The PANAS consists of twenty items, comprising of two ten-item subscales assessing positive affect (interested, excited, enthusiastic, strong, alert, inspired, determined, attentive, active, proud) and negative affect (upset, distressed, guilty, scared, hostile, irritable, ashamed, nervous, jittery, afraid). Participants were asked to rate the extent to which they felt each emotion both before and after their VR experience using a five-point Likert scale ranging from 1 (“very slightly or not at all”) to 5 (“extremely”). The PANAS has demonstrated reliability and construct validity across psychological research (Crawford et al., 2004) and has also been employed in studies examining emotional responses to art experiences (see for example Drake et al., 2024; Holt, 2018). Due to a procedural oversight, the negative affect item *afraid* was inadvertently omitted from the initial version of the questionnaire. The omission was corrected partway through data collection; however, as a result, responses for this item are limited and should be interpreted with caution due to the reduced sample size. This has been highlighted within the results section. A full list of questions can be found in the appendix.

Qualitative data

The questionnaire included open-ended questions designed to elicit qualitative insights into the participants’ emotional responses and to gather feedback on their experiences with and interactions across the two VR interfaces. A summative content analysis approach was used to analyse the qualitative data (Hsieh & Shannon, 2005). This method was selected to enable systematic coding and categorisation within responses to identify recurring themes, patterns, and points of divergence between participant groups and VR experiences.

Responses to the survey question: ‘*Would you suggest additional features to help us improve the interface?’* were initially read and annotated by the research team to develop familiarity with the data and develop an initial coding framework. Preliminary thematic categories were elaborated upon collaboratively during a design workshop held in January 2025 with key project partners. The group collectively reviewed a sample of responses and co-developed a coding schema grounded in both prior research on affect and voice interaction and emerging patterns in the data. This collaborative analysis ensured that the coding framework reflected a range of disciplinary perspectives, from digital art to psychology, computer science and gaming.

Following the workshop, for the other questions responses were systematically coded using this schema, and themes were refined through iterative discussion between this paper’s authors. The content meaning of sentences or paragraphs within the free text responses transcribed were labelled. Code frequency was tallied independently between the two authors and subsequently compared. Any outliers were discussed before a consensus on coding was reached. Comparisons were then made across participant groups and experiences to explore how different users interacted with and interpreted the interfaces. Having an artist and psychologist work together was valuable in ensuring participants' responses were adequately interpreted. The artist approached this from the viewpoint of designing the experience and having a particular expectation and intention of the participant's experience, while the psychologist had the lens of previous experience in measuring wellbeing in relation to short artistic experiences.

RESULTS

Quantitative findings

Quantitative data were analysed using IBM SPSS Statistics (Version 30). Repeated-measures ANOVA was conducted to examine the effects of the interface (*VRoar* vs. *AudioVirtualizer)* and group (singer vs. non-singer) on PANAS scores. Within-subject contrasts and interaction effects were explored for significant findings. Where relevant, Bonferroni corrections were applied to control for Type I errors across multiple comparisons.

Follow-up paired-sample t-tests were conducted for each PANAS item to examine which specific emotions showed changes over time. Effect sizes were calculated using Cohen's d, with Hedges' g reported to adjust for small sample bias. Significance thresholds were adjusted based on the number of comparisons.

A total of 91 participants took part in the study. Five were removed due to incomplete datasets (defined as more than three missing PANAS scale items),[1] leaving a final sample of 86 participants. Of these, 40 participants identified as singers and 46 as non-singers. Within the singer group, the level of vocal experience ranged from beginner to professional, with 65% identifying as either semi-professional or professional. Participants ranged in age from 18 to over 65 years, with the largest age group being 18-24 years (36%). In terms of gender identity, 57% identified as female, 38% as male, 4% as non-binary, and 2% as 'other'. Computer game use varied across the sample: 27% of participants reported they *never* played, 30% played *once a month or less*, 15% played *once a week*, and 28% played *daily or almost daily*. The majority of participants (72%) reported having previous experience with VR technology.

Participants were relatively balanced across gender, age, and gaming experience between singer and non-singer groups and between *VRoar* and *AudioVirtualizer* conditions. The only significant difference was that singers were more likely to have prior VR experience than non-singers, $\chi^2(1, N = 86) = 10.86$, $p < .001$.

Positive/Negative Affect Subscales

A 2 (Time: Pre, Post) × 2 (Scale: Positive Affect, Negative Affect) × 2 (Singer: Singer, Non-Singer) × 2 (VR Experience: *VRoar*, *AudioVirtualizer*) mixed-design ANOVA was conducted to examine the effects of singing group and VR experience on affective outcomes over time. Figure 6 displays the individual graphs for positive/negative affect subscales for each group and VR experience.

There was a significant main effect of Time, $F(1, 82) = 9.94$, $p = .002$, $\eta^2_p = .108$, indicating a change in affect scores from pre- to post-experience. There was also a significant main effect of subscale (positive vs. negative affect), $F(1, 82) = 585.78$, $p < .001$, $\eta^2_p = .877$, suggesting a difference in overall scores between positive and negative affect.

A significant Time × Emotion Subscale interaction was observed, $F(1, 82) = 22.98$, $p < .001$, $\eta^2_p = .219$, indicating that the change in affect over time differed depending on whether the scale was positive or negative. No significant Time × Subscale × Group interaction (Singers/Non-Singers) was found, $F(1, 82) = 0.76$, $p = .387$, $\eta^2_p = .009$, suggesting that the pattern of affect change over time did not significantly differ

1 Please note that some of these participants did provide open ended responses to the broader questions asked in the second part of the questionnaire. These responses were included within the qualitative analysis.

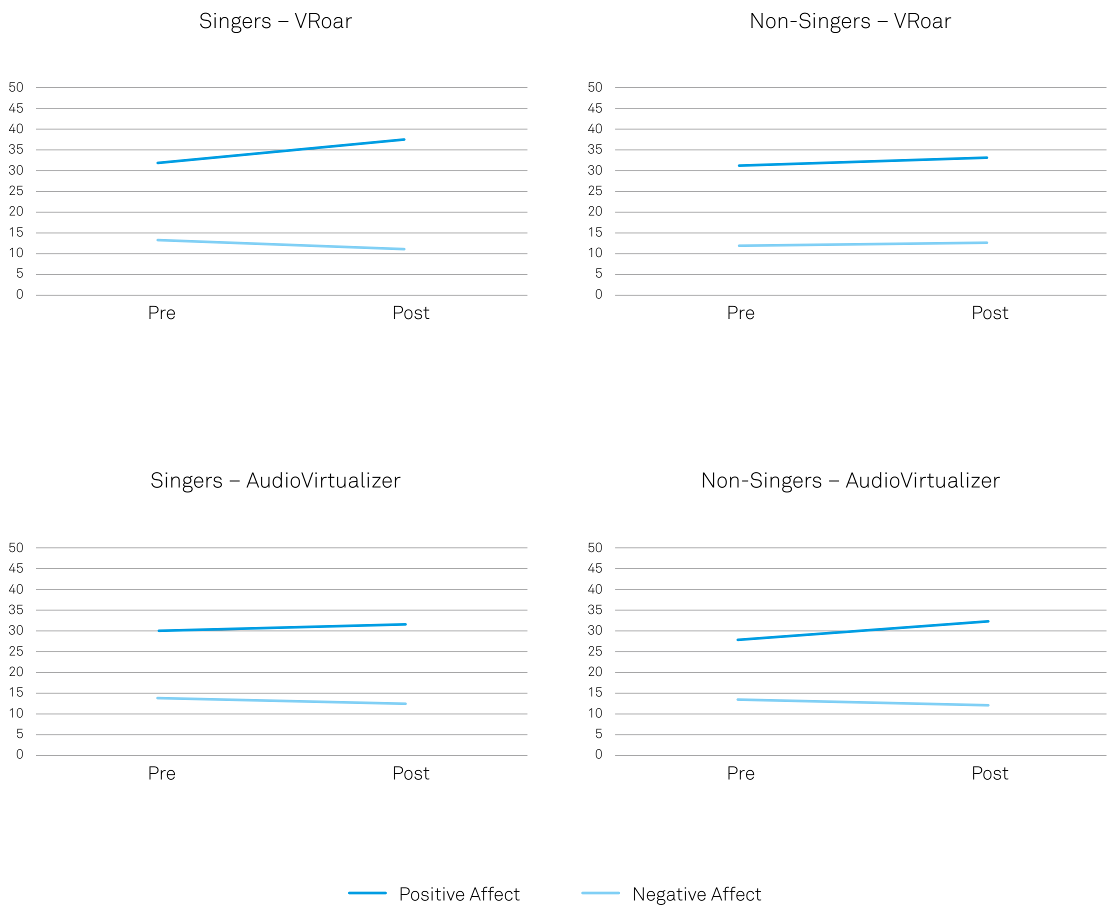


FIGURE 6. Mean positive and negative affect scores before and after the VR experience, presented separately for singers and non-singers across *VRoar* and *AudioVirtualizer* conditions.

between singers and non-singers. Similarly, the Time × Subscale × VR Experience interaction was not significant, $F(1, 82) = 0.17$, $p = .685$, $\eta^2_p = .002$. However, a significant Time × Scale × Group × VR Experience interaction emerged, $F(1, 82) = 6.98$, $p = .010$, $\eta^2_p = .078$. This indicates that the combined effects of group and VR experience influenced how positive and negative affect changed over time.

To investigate the significant Time x Emotion Subscale x Group x VR experience interaction, follow up within-subjects contrasts were conducted for each subgroup defined by VR experience (*VRoar* vs. *AudioVirtualizer*) and group (singer vs. non-singer).

VRoar Experience

Among participants in the *VRoar* condition, a significant Time × Scale interaction was found for singers, $F(1, 19) = 14.67$, $p = .001$, $\eta^2_p = .436$, indicating that changes in affect from pre- to post-experience differed between the positive and negative affect scales. This was accompanied by a strong main effect of time, $F(1, 19) = 286.40$, $p < .001$, $\eta^2_p = .938$, and a smaller but significant main effect of scale, $F(1, 19) = 4.90$, $p = .039$, $\eta^2_p = .205$. In contrast, for non-singers in the *VRoar* group, the Time × Scale interaction was not significant, $F(1, 19) = 0.50$, $p = .488$, $\eta^2_p = .026$, suggesting that both positive and negative affect changed similarly over time. Nonetheless, significant main effects of time, $F(1, 19) = 101.38$, $p < .001$, $\eta^2_p = .842$, and scale, $F(1, 19) = 4.74$, $p = .042$, $\eta^2_p = .200$, were observed.

AudioVirtualizer Experience

For participants in the *AudioVirtualizer* condition, no significant Time (Before/After) × Subscale (Pos/Negative Emotion) interaction was found among singers, $F(1, 19) = 1.92$, $p = .182$, $\eta^2_p = .092$, indicating similar patterns of change for both affect scales. A significant main effect of time was observed, $F(1, 19) = 114.70$, $p < .001$, $\eta^2_p = .858$, but there was no main effect of scale, $F(1, 19) = 0.00$, $p = 1.000$. In contrast, non-singers in the *AudioVirtualizer* group demonstrated a significant Time × Subscale interaction, $F(1, 25) = 16.68$, $p < .001$, $\eta^2_p = .400$, suggesting a differential pattern of change in positive and negative affect. This was supported by significant main effects of time, $F(1, 25) = 154.23$, $p < .001$, $\eta^2_p = .861$, and scale, $F(1, 25) = 6.20$, $p = .020$, $\eta^2_p = .199$.

Individual emotions

An item-level follow-up analysis was also conducted to understand whether there were any specific positive or negative emotional changes in participants after experiencing *VRoar* or *AudioVirtualizer*. A Bonferroni correction was applied to control for multiple comparisons, setting the adjusted alpha level at $\alpha = .0025$. Bonferroni corrections were applied separately within each subgroup (e.g., singer vs. non-singer, *VRoar* vs. *AudioVirtualizer*), resulting in a threshold of $p < .0025$ (0.05/20) for each table of PANAS items. This approach was chosen to balance Type I and Type II error risks, reflecting the conceptual independence of the subgroups rather than applying a single conservative correction across all 80 comparisons (which would yield $p < .000625$). Five PANAS items showed statistically significant changes from pre- to post-experience after applying a Bonferroni correction ($\alpha = .0025$) for singers in the ***VRoar*** experience.

- Inspired significantly increased, $t(19) = -5.40$, $p < .001$
- Alert significantly increased, $t(19) = -4.33$, $p < .001$
- Active significantly increased, $t(18) = -3.53$, $p = .001$
- Irritable significantly decreased, $t(19) = 3.20$, $p = .002$
- Nervous significantly decreased, $t(19) = 3.29$, $p = .002$

These results indicate that singers exposed to the *VRoar* experience experienced marked increases in high-arousal positive emotions (*Inspired*, *Alert*, *Active*), along with reductions in negative affective states linked to agitation and anxiety (*Irritable*, *Nervous*). The pattern suggests that *VRoar* may be particularly effective in enhancing emotional engagement and reducing emotional tension for individuals with a vocal or musical background.

In the *VRoar* non-singer group, only one PANAS item showed a statistically significant change following Bonferroni correction (α = .0025). *Alert* significantly increased from pre- to post-experience, $t(19) = -4.41, p < .001$. This result suggests that the *VRoar* condition enhanced alertness in non-singers, potentially indicating a heightened state of focus or engagement following the immersive experience.
No other items reached significance after correction.

No PANAS items reached statistical significance at the Bonferroni-adjusted threshold for singers experiencing ***AudioVirtualizer***. For non-singers, four PANAS items showed statistically significant increases after Bonferroni correction (α = .0025):

- Interested, $t(25) = -3.09, p = .002$
- *Excited*, $t(24) = -4.24, p < .001$
- *Enthusiastic*, $t(25) = -3.11, p = .002$
- *Inspired*, $t(24) = -4.63, p < .001$

There were no significant changes observed in negative affect scale items. These findings suggest that the *AudioVirtualizer* condition elicited strong increases in high-arousal, positively valanced emotional states among non-singers, including feelings of energy, motivation, and interest. No negative affect items met the Bonferroni-corrected threshold. Across both interfaces there was a less noticeable change in the negative emotions overall, which unsurprising as most people would have started the VR experience already feeling fairly positive. The fact that we had less data for the item afraid is unlikely to have changed the results significantly, as it was not an emotion we would have expected to change significantly after having done the experience.

Qualitative findings

Open-ended questions were analysed using summative content analysis (to enable the systematic coding and categorisation of data into themes, patterns and points of divergence between singers and non-singers. Responses were initially read and annotated by the research team to become familiar with the data and develop an initial coding framework. This collaborative analysis ensured that the coding framework reflected a range of disciplinary perspectives. Following this initial group coding exercise, themes were further refined for each question, through iterative discussion between this paper's authors. Comparisons were made across participant groups and experiences to explore how different users interacted with and interpreted the interfaces.

Overall, the qualitative findings mirrored the quantitative findings but gave much more detailed information regarding which features were most appreciated in the different interfaces, and were compared between the two groups. A detailed description of the qualitative questions, codes and percentages are included in Tables 1-4. The following section presents findings relating to two elements of experience: 1) wellbeing and 2) interface experiences and suggested features.

Wellbeing

In response to the question ***'how did you feel emotionally after experiencing this interface, as opposed to before?'*** participants reported distinct affective responses that appeared to be shaped by the design of each interface.

AudioVirtualizer was associated with more frequent reports of positive affect, these were commonly framed in terms of calmness or stress relief. For instance, some non-singers described feeling "*more relaxed as the experience was an opportunity to take my mind off work-related worries*" (participant 37) or "*calm… pulled away from my day-to-day routine*" (participant 38). *VRoar* also elicited a high frequency of positive affective responses which seemed to be more strongly worded. For example, some described feeling "*uplifted, inspired,* *happy, grounded, focused, optimistic*" (participant 53) and "*really giddy… very positive and happy*" (participant 50). These examples illustrate broader trends in how the interfaces shaped emotional responses.

Curiosity/play and vocal or sound exploration were commonly reported in *VRoar* and *AudioVirtualizer*, although vocal or sound exploration was not coded amongst non-singers in the *AudioVirtualizer* condition for this question. Some singers using *VRoar* described the experience as creatively freeing, with one participant noting the opportunity to "*explore all the extremities of [their] voice… playful and accessible without judgement*" (participant 69). In contrast, non-singers using *AudioVirtualizer* sometimes expressed curiosity alongside a desire for more guidance or time to explore, with illustrative comments such as a wish to explore more "*with guidance*" (participant 23).

Negative affect was reported more frequently in *AudioVirtualizer* with some non-singers describing feelings of disorientation or sensory overload – e.g., feeling "*claustrophobic at times*" (participant 42) or the experience being "*draining*" (participant 30). In contrast, *VRoar* elicited fewer reports of negative affect. When mentioned, these responses were typically less intense and often described by non-singers, such as feeling "*spaced out*" or "*on edge*" (participant 26) or indicating a positive experience, such as when they said they were "*sad to leave*" (participant 56). Overall, these findings suggest that while both interfaces were generally experienced as emotionally engaging, *VRoar* may have more effectively supported creative exploration and emotional uplift, particularly for those with prior vocal experience.

This is compiled in Table 1 below.

Participants were also asked ***'Did you feel inspired to make different sounds?'*** (which links directly to the 'inspired' emotion in the PANAS Scale). A total of 74% of *VRoar* participants reported feeling inspired by the experience, compared to 61% in the *AudioVirtualizer* condition, indicating that *VRoar* elicited a greater sense of inspiration from participants. This was especially the case for the singers, where 86% answered positively. This supports the quantitative results which also suggested that *VRoar* is most engaging for those with vocal experience.

As a follow up, participants were asked ***'If you felt inspired to make new sounds, please try to explain how you feel in your own words.'***

In *AudioVirtualizer,* what most inspired both singers and non-singers alike was the ability to see the sound reflected in the visuals, though singers more often mentioned the role of their own voice in this. Positive affect was the second most frequently coded item across both groups. Many participants described how experimenting with vocal sounds (from humming to singing, to making unusual noises) produced dynamic visual feedback that enhanced their sense of agency and enjoyment. The interface encouraged play and exploration equally amongst both groups, but non-singers more frequently said the visual elements caused a feeling of inspiration. Singers also commented on the novelty of the experience whereas non-singers did not.

For *VRoar,* the most frequently coded responses related to 'curiosity, play and exploration.' Participants reported a wide range of vocal play to explore the VR environment, from pitch shifts and vibrato to rhythmic pulses and spontaneous utterances. This, as with *AudioVirtualizer,* was framed as an engaging and enjoyable experience. Positive affect was almost equally frequently coded with singers and non-singers reporting feelings of excitement, freedom, and creative satisfaction. This suggests that *VRoar* effectively supported intuitive exploration and engagement, which in turn, could have contributed to participants' reporting of feeling inspired.

Interface experiences and suggested features

When participants were asked ***what they liked most about the interface***, for *AudioVirtualizer* this was most often the visual

		AV-NS	%	AV-S	%	VR-NS	%	VR-S	%
Total responses		27		20		19		22	
1. Positive affect	Clear expressions of positive emotional responses such as enjoyment, happiness, excitement, satisfaction, fun, or uplifted mood.	23	85.2 %	14	70%	14	73.7 %	17	77.3 %
2. Negative affect	Clear expressions of unpleasant emotional states such as frustration, anxiety, confusion, disappointment, disorientation, experiencing discomfort or feeling overwhelmed.	10	37 %	7	35 %	3	15.8 %	2	9.1 %
3. Neutral/no change in affect	Explicit statements that no emotional change occurred	2	7.4 %	1	5%	3	15.8 %	4	18.1 %
4. Vocal/sound experimentation to see effect	Mentions trying out different vocalisations (pitch, rhythm, volume, words) to observe effects or responses in the environment	0		0		5	26.3 %	3	13.6%
5. Curiosity/ play/ exploration	Reflects exploratory or playful engagement with the experience, often driven by a desire to see what might happen.	3	11.1 %	6	30%	6	31.6 %	6	27.3%

TABLE 1. Responses to the question:
'How did you feel emotionally after experiencing this interface, as opposed to before?'

		AV-NS	AV-NS	AV-S	AV-S	VR-NS	VR-NS	VR-S	VR-S
		18		14		12		23	
			per-centage		per-centage		per-centage		per-centage
1. Positive affect	Clear expressions of positive emotional responses such as enjoyment, happiness, excitement, satisfaction, fun, or uplifted mood.	11	61.1 %	8	57.1%	7	58.3%	11	47.8%
2. Negative affect	Clear expressions of unpleasant emotional states such as frustration, anxiety, confusion, disappointment, disorientation, experiencing discomfort or feeling overwhelmed.	2	11.1%	3	21.4%	1	8.3%	0	
3. Neutral/ no change in affect	Explicit statements that no emotional change occurred.	0		0		0		1	4.4
4. Vocal/sound experimenta-tion to see effect	Mentions trying out different vocalisations (pitch, rhythm, volume, words) to observe effects or responses in the environment.	17	94.4%	10	71.4%	9	75%	19	82.6%
5. Curiosity/ play/ exploration	Reflects exploratory or playful engage-ment with the experi-ence, often driven by a desire to see what might happen.	10	55.6 %	8	57.1%	12	100%	15	65.2%
6. Instructions and guidance	Comments on the need for, or value of, clearer instructions, support or orientation within the experience.	1	5.6%	2	14.3%	0		1	4.4%
7. Interface and navigation	Ease of use and movement through the VR experience, in-cluding controls and scene transitions.	2	11.1%	0		0		0	
8. Visual and aesthetic features	Refers to visual elements such as colours, textures, shapes, or the overall look and feel of the interface.	10	55.6%	2	14.3%	2	16.7%	2	8.7%
9. Novelty	Highlights the new-ness or uniqueness of the experience.	0		2	14.3%	0		0	

TABLE 2. Coded responses to the question:
'*If you felt inspired to make new sounds, please try to explain how you feel in your own words.*'

	AV-NS	AV-NS	AV-S	AV-S	VR-NS	VR-NS	VR-S	VR-S
Total responses	27		20		18		21	
	number	per-centage	number	per-centage	number	per-centage	number	per-centage
1.Voice-responsive interaction/audio-based interaction	7	26%	6	30%	11	61%	11	52.38%
2. Mapping choices	4	15%	0	0%	4	22%	2	9.52%
3. Aesthetics of models or backdrops	20	72%	9	45%	5	27.78%	6	28.57%
4. The immersion it provided	7	26%	2	10%	3	16.67%	5	23.81%
5. The fun, entertainment or play it provided	3	11%	5	25%	3	16.67%	8	38.09%
6. The ability to explore	3	11%	2	10%	3	16.67%	2	9.52%
7. Feeling of Freedom	0	0%	0	0%	0	0%	4	19.05%
8. It provided relaxation	1	3.70%	0	0%	0	0%	0	0%
9. Music, different sound options through the menu	4	15%	2	10%	0	0%	0	0%
10. It's a new, novel and or experimental experience.	3	11%	5	25%	0	0%	2	9.52%

TABLE 3. Responses to the question:
'What did you like most about the interface (AudioVirtualizer/VRoar)?'

aesthetics or the music in the interface, whereas for *VRoar* it was the voice-responsive interaction. In *AudioVirtualizer* across both groups, voice-responsive interaction came second, followed by the fun (singers) or immersion (non-singers) it provided. The novelty of the experience was more often commented upon by singers experiencing *AudioVirtualizer* than singers experiencing *VRoar*. *VRoar* was enjoyed in subtly different ways by singers and non-singers. Singers had more fun than non-trained vocalists and frequently mentioned a feeling of freedom, whereas non-singers appreciated the aesthetics and mappings more often.

When asked, ***'Would you suggest additional features to help us improve the interface?'*** both singers and non-singers for *AudioVirtualizer* requested more guidance, either prior to the experience or within the interface. Singers requested mostly in-interface guidance, such as '*a little person to guide you'* (participant 64), instructions on what the buttons do, or a menu with the instructions or using visual highlights as a means to show where objects could be selected. Non-singers requested both more verbal instruction prior to the experience and in-interface instructions like prompts about ways to use the voice, and a demo tutorial showing what different controls can do. Non-singers also reported challenges with operating the controls and navigating in the worlds. In *VRoar*, there were fewer suggestions relating to guidance, but when it was requested by non-singers, this was related to instructions to better understand 'the task,' whereas singers only asked for some basic instructions or tips to get the most out of the interface.

For *VRoar,* the most frequent theme was 'mappings' across both groups, but there were specific requests from singers and non-singers. Singers wanted pitch and timbre changes to be more obviously represented in the graphics, and for the balls to not just move but change in size or shape. One singer also requested more response to unpitched sounds like plosives and fricatives, with the link to voice patients who may not be able to fully phonate. Non-singers asked for the colours of the balls to change more, for different objects to respond differently to different sounds to encourage more sound experimentation, with one participant suggesting that objects disintegrate when certain pitches were reached. A noted difference here is that singers more frequently asked for more variety in the sound features picked up and non-singers wanted more variation in the graphic responses.

Across both interfaces, participants wanted more options. For *AudioVirtualizer,* non-singers wanted more levels and more activities to do. Singers wanted more variety, more animation and a different way to change rooms. For *VRoar,* one singer suggested polyphonic possibilities, and another wanted to collaborate with someone else in the same digital space. Non-singers wanted more interaction options. One suggested different objects on different levels, and another wanted to directly manipulate the mapping algorithm.

DISCUSSION

This study sought to understand whether VR art experiences that are voice-responsive have an impact on self-reported positive and negative affect, with a particular focus on the differences between singers and non-singers. There were interesting differences observed between the two groups, with singers showing a significant positive change in mood for *VRoar* and non-singers showing a significant positive change in emotions when experiencing the *AudioVirtualizer*. There are myriad reasons why these differences were observed across groups. For example**,** non-singers from our sample group had a higher need to receive guidance, instructions and prompts. This may be partially because our sample of singers was more likely to have prior VR experience than non-singers, but could also be due to singers being more comfortable experimenting with their voice to the extent that it was linked to a sense of 'freedom.' Having previous VR experience is unlikely to affect how comfortable someone is to use their voice in an expressive manner, so

Experience group		AV-NS	AV-NS	AV-S	AV-S	VR-NS	VR-NS	VR-S	VR-S
Total Responses		25		18		16		16	
		number	per-centage	number	per-centage	number	per-centage	number	percent-age
1. Interaction	This concerns the interaction possibilities the interface provides.	5	20%	3	16.7%	4	25%	6	37.5
2. Options	When people asked for more or different options in the interface this code was used.	7	28%	4	22,2%	8	50%	5	31.30%
3. Challenge	This was coded when people mentioned something was too difficult or if they felt a challenge was lacking and should be added.	8	32%	3	16.7%	2	12.50%	1	6.25%
4. Aesthetics	This has to do with the experience of the imagery or the music and sound and how this effects the user.	3	12%	5	27.8%	4	25%	5	31.30%
5. Guidance	This concerns how much verbal or in-interface instruction the user is getting.	8	32%	5	27.8%	4	25%	2	12.50%
6. Mappings	This relates to the computer coding between input and output, i.e. in this case: how does a change in a sound feature affect the imagery.	2	8%	2	11.1%	8	50%	7	43.80%
7. Latency	How fast is the visual change noticeable after modulating the sound.	0		0		0		3	18.8

TABLE 4. Responses to the question:
'***Would you suggest additional features to help us improve the interface?***'

the request for prompts to use the voice amongst the non-singer group was most likely due to a lack of confidence in using their voice. However, request for guidance in using the controls could be to do with a lack of previous VR experience. These two different needs should be slightly differently addressed. The need for ideas on how to vocalise could be an audio feature where suggestions are made vocally to keep the user engaged with the interface – a mode that could be especially helpful for non-singers. The need to be reminded of how the controls work could be a visual overlay that could be activated upon a button switch, thus making it optional and useable by both groups.

Singers also more frequently commented on the novelty of the experience, which may indicate that in this group there is a bigger need for interfaces like this to exist. As we aim to increase wellbeing amongst both groups, it is important that we design to meet the need for more guidance amongst non-singers, whilst keeping the interface uncluttered for singers who were happy to experiment without further instructions that could constrain their experience. Our findings therefore illustrate the complexity of the different dimensions of affect: rather than a uniform 'feel-good' effect, the experiences shaped emotions differently depending on users' prior skills, comfort with vocal expression, and familiarity with VR technology. Although our findings focus on short-term emotional changes, these momentary increases in positive affect may be understood as part of the micro-processes that underpin wellbeing, building on the idea that experiences held 'in the moment' can provide meaningful emotional and relational value within arts and participatory contexts (e.g. Holt, 2020; Dowlen et al., 2022). Within therapeutic and participatory art contexts, such shifts can create openings for self-expression, stress release, and reconnection with one's own body and voice. The experience of curiosity and play described by participants, for example, resonates with models of wellbeing that emphasise engagement and creativity as restorative processes (e.g. Fredrickson, 2001). In this sense, the artworks may function as rehearsal spaces, supporting participants in exploring and regulating emotions in a safe and embodied way.

One explanation for the differences between singers and non-singers lies in the ways participants approached the vocal interaction, with singers tending to describe the experience as 'freeing' and engaging readily in vocal experimentation. This aligns with Fredrickson's (2001, 2004) broaden-and-build theory, which suggests that positive emotions expand people's thought-action repertoires, making them more willing to play, explore and create. In contrast, non-singers more frequently requested guidance, prompts or structure, suggesting a higher baseline of uncertainty or self-consciousness. For this group, supportive clues within the interface may be important for reducing negative affect, creating a feeling of a 'safe space' where experimentation feels less intimidating.

The novelty of the experience also seems to play a role, as singers were more likely to comment on the uniqueness of interacting through their voice in VR, which may have heightened their sense of engagement and contributed to an upward shift in mood. Non-trained singers may have initially felt more constrained by the unfamiliarity of both VR and vocal improvisation, requiring scaffolding to reach a similar level of emotional benefit. It is important to note from this that positive affect in VR experiences may not emerge automatically but depends on balancing novelty with accessibility. Too much openness may overwhelm those with less vocal confidence, while too much structure may limit the exploratory benefits of more confident users. This has implications from a design perspective for the next iterations of this project. It is necessary to consider the role of guidance in supporting the confidence of non-vocalists while also ensuring those with vocal experience preserve a sense of freedom and exploration that fuels their positive emotional response.

The voice-responsive design also holds promise for accessibility and inclusion. For participants with limited mobility or for whom traditional movement-based interaction is challenging, voice input offers a low-barrier entry point into creative VR experiences. This inclusive design aligns with wellbeing frameworks that emphasise autonomy, participation, and sense of agency as key contributors to psychological health (e.g. Putney et al., 2024).

The navigation issues in *AudioVirtualizer* were mostly resolved in *VRoar.* Hence, the navigation method in *VRoar,* which relies on the joystick for left, right, forwards and backwards and direction of gaze with forwards or backwards for up or down in height, will be copied over in the next interface. Furthermore, it seems that the principle of different 'rooms' from *AudioVirtualizer* should be brought back in the next piece as people across both groups seemed to enjoy that variety and a few people experiencing the *VRoar* interface commented on finding it confusing that only the background and mappings changed but the landscape stayed the same when pressing the 'A' button. As the next system will be built upon detecting and reflecting different emotions from voice and language analysis, this could be the method by which someone reaches a different 'room' or 'level:' through expressing different emotions vocally and verbally.

Another interesting thing to notice was that more people mentioned aesthetics as the best thing about the interface for the *AudioVirtualizer* (especially non-singers), whereas for *VRoar* the voice responsiveness category was the most popular. Due to there being three distinctly different rooms in the *AudioVirtualizer,* with new objects and artwork in each environment, there was more variation in the *AudioVirtualizer*; this would be important to consider and include in future work.

Further comments that did not fall specifically under a tagged category but that are worth considering moving forward was that some people when experiencing certain scenes in the *AudioVirtualizer* felt uncomfortable having nowhere to stand on. As other people said that seeing the shapes disappear underneath them increased the experienced immersion, a balance could be to have a ridge or stepping stones to walk on with the abyss appearing underneath the user.

CONCLUSIONS AND FURTHER WORK

Through the evaluation of two interfaces, we learned that a short-term increase in positive emotions is achievable from designed VR art experiences. *AudioVirtualizer* achieved this for the group of non-singers and the *VRoar* interface achieved it for singers. Singers and non-singers engaged differently, suggesting a tailored design is required. Prior to the VR experience, a more extensive verbal briefing should be given to non-singers and those with less VR experience. This could be further strengthened with in-interface prompts to stimulate people to experiment with their voice. For non-singers, it may help to bring the option of having precomposed music control the graphics (like in *AudioVirtualizer*) back in rather than relying entirely on voice control. To summarise, the results imply two principles for future voice VR art:

(1) providing guidance and optional support modes for novices to reduce anxiety, and
(2) maximising flexibility (i.e. the lack of specific instructions inside the interface) and novelty for experienced vocalists to sustain engagement.

A more detailed analysis of which aesthetic features increase positive emotion will be necessary to obtain a clearer picture on exactly how a VR interface can improve emotional wellbeing for the two different groups, though non-singers may benefit from an in-interface tutorial that goes over the functionality of the controls and explains the mappings to help them gain confidence. To further inspire the singers, further variety in the types of vocal utterances that are reflected in the mappings should be added, for example through allowing the interface to detect the difference between plosives and fricatives, vowels and consonants, and voice quality. Another way we plan to make the next interface more varied is by making the interface capable of distinguishing the emotional difference in vocal utterances and sentences and have this mapped to emotionally expressive animated graphics. We plan to design 3D models expressive of four emotions – happy, sad, angry and calm, each residing in one of the four quadrants of Russell's circumplex model of affect (Russel, 1980) – and animating each shape according to the same four emotions. This way, a happy shape could be animated in a happy, sad, calm or angry way, allowing for multiple mappings and the presence of ambiguity, the same way that we can use happy words but say them in a sad way.

While these results cannot yet confirm long-term wellbeing benefits, they highlight how creative VR experiences can facilitate immediate emotional uplift and self-expression, which are both integral components of subjective wellbeing. The combination of aesthetic engagement and embodied voice interaction presents a promising direction for future research in VR arts and wellbeing, particularly in exploring how these digital spaces might support emotional awareness, agency and connectedness. The results of this study suggest voice-responsive VR art could be leveraged in therapeutic or inclusive art programmes to enable engagement and wellbeing.

ACKNOWLEDGMENTS

This research was made possible by an AHRC Catalyst grant, grant nr: AH/Z506618/1

VRoar was created with Arts Council England funding. Thanks to Pippa Anderson for undertaking the evaluations in the Northeast of England and to Carlo Harvey and Nick Collins for feedback on this paper.

Creative UK funding facilitated a small-scale evaluation of *AudioVirtualizer* in 2023.

REFERENCES

Allison, Fraser, Carter, Marcus and Gibbs, Martin (2020) 'Word Play: A History of Voice Interaction in Digital Games', Games and Culture, Volume 15, Issue 2, March 2020, Pages 91-113, https://doi.org/10.1177/1555412017746305.

Bezan, Sarah (2022) 'The Species Revivalist Sublime Encountering the Kaua'i 'Ō'ō Bird in Jakob Kudsk Steensen's Re-Animated' – Valérie Bienvenue and Nicholas Chare (eds), Animals, Plants and Afterimages: The Art and Science of Representing Extinction, Brooklyn NY: Berghahn Books, 211–227.

BMFVR Gaming, 2022, *FREE Quest 2 Harry Potter VR Game Has Voice Spell Casting!, Screencast, available at:* https://www.youtube.com/watch?v=GL3-j4dWTaw, Accessed 14/11/25.

Butler, Julie and Kern, Margaret L. 2016.The perma-profiler: A brief multidimensional measure of flourishing, International Journal of Wellbeing. 6, 1–48 (2016). https://doi.org/10.5502/ijw.v6i3.526.

Clift, Stephen 2012. Singing, Wellbeing, and Health. In: Music, Health and Wellbeing, edited by Raymond MacDonald, pp. 113–124. Oxford: Oxford University Press.

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. British journal of clinical psychology, 43(3), 245–265.

Davies, Charlotte A. (2005), Landscapes of ephemeral embrace: a painter's exploration of immersive virtual space as a medium for transforming perception, Thesis, University of Plymouth, Available at: http://dx.doi.org/10.24382/4437.

Deshmukh, Akshay Madhav, and Chalmeta, Ricardo 2024. "User Experience and Usability of Voice User Interfaces: A Systematic Literature Review" *Information* 15, no. 9: 579. https://doi.org/10.3390/info15090579.

Dowlen, Robyn, Keady, John, Milligan, Christine, Swarbrick, Caroline, Ponsillo, Nick, Geddes, Lucy, and Riley, Bob (2022) In the moment with music: An exploration of the embodied and sensory experiences of people living with dementia during improvised music-making. Ageing & Society, 42(11), 2642–2664.

Drake, J. E., Eizayaga, M., & Wawrzynski, S. (2024). Making versus viewing art: Effects on affect, enjoyment, and flow. Empirical Studies of the Arts, 42(1), 147–165.

Dunphy, Bryan (2022), *Sculpting Unrealities: Using Machine Learning to Control Audiovisual Compositions in Virtual Reality*. Doctoral thesis, Goldsmiths, University of London.

Edgar, Amanda N. 2019. Culturally speaking: the rhetoric of voice and identity in a mediated culture. Columbus: The Ohio State University Press.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. American psychologist, 56(3), 218.

Fredrickson B.L. (2004) The broaden-and-build theory of positive emotions. Philos Trans R Soc Lond B Biol Sci. 2004 Sep 29;359(1449):1367-78. doi: 10.1098/rstb.2004.1512. PMID: 15347528; PMCID: PMC1693418.
Gosling, E. (2018) Marshmallow Laser Feast's Colossal Wave, a stunning AV piece that shows how people are messing up the world's oceans, Creative Boom, Available at: <https://www.creativeboom.com/inspiration/marshmallow-laser-feast/>, accessed 10/11/25.
Harma, Vuokko 2011. Interaction and Performativity in Digital Art Exhibitions. Norsdisk Museology. 1. S. 98–105. https://journals.uio.no/museolog/article/view/3148/2724.
Holt, N. J. (2018). Using the experience-sampling method to examine the psychological mechanisms by which participatory art improves wellbeing. Perspectives in public health, 138(1), 55–65.
Holt, N.J. (2020) Tracking momentary experience in the evaluation of arts-on-prescription services: using mood changes during art workshops to predict global wellbeing change. Perspectives in public health, 140(5), 270–276).
Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. Qualitative Health Research, 15(9), 1277–1288. https://doi.org/10.1177/1049732305276687.
Jean-Berluche, Ducel (2024) 'Creative expression and mental health', Journal of Creativity, 34, 2, 1-6, Available at: https://doi.org/10.1016/j.yjoc.2024.100083.
Kang, Jing, Scholp, Austin and Jiang, Jack J. 2018. A Review of the Physiological Effects and Mechanisms of Singing. Journal of Voice. 32. 4. 390–395. https://www.jvoice.org/article/S0892-1997(17)30223-0/abstract.
van 't Klooster, Adinda and Collins, Nick 2021, 'Virtual Reality and Audiovisual Experience in the AudioVirtualizer', the Journal EAI Endorsed Transactions on Creative Technologies, 8., 27 (March 2021). 1–7. https://eudl.eu/doi/10.4108/eai.19-3-2021.169037.
van 't Klooster, Adinda and Collins, Nick (2024) 'Voice Responsive Virtual Reality' In Proceedings New Interfaces for Musical Expression. NIME2024, Utrecht, https://nime.org/proceedings/2024/nime2024_51.pdf.
Putney, Henry, Siliver, Sarah, Silvia, Paul, Christensen, Alexander P, and Cotter, Katherine N (2025) 'Why does creativity foster wellbeing? Autonomy, competence, and relatedness during everyday creative activities,' Journal of Research in Personality, 113. https://doi.org/10.1016/j.jrp.2024.104552.
Pasquier, Phillippe (2018) 'Respire: a Breath Away from the Experience in Virtual Environment', CHI'18 Extended Abstracts, April 21–26, 2018, Montreal, QC, Canada.
Russell, James, A. (1980) 'A Circumplex Model of Affect', Journal of Personality and Social Psychology 39 (1980) pp. 1161–1178.
SOSO Limited, 2003-2025, Miraj, Available at: <https://www.sosolimited.com/work/miraj>, Accessed 13/11/25.
SOSO Limited, 2016, *Making an interactive artwork that you paint with your voice, Available at:* <https://medium.com/sosolimited/making-an-interactive-artwork-that-you-paint-with-your-voice-c1a8b0e033f5>, Accessed 13/11/25.
Steensen, Jacob, K. (2025), Jakob Kudsk Steensen, Available at: <https://www.jakobsteensen.com/the-deep-listener>
Watson, David and Clark, Lee A. 1988. 'Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales,' Journal of Personality and Social Psychology, 54. 6. 1063–1070.
Wagener, Nadine, Schöning, Johannes, Rogers, Yvonne, and Niess, Jasmin (2021) 'The Role of Mobile and Virtual Reality Applications to Support Well-Being: An Expert View and Systematic App Review' In: Ardito, C and Lanzilotti, R and Malizia, A and Petrie, H and Piccinno, A and Desolda, G and Inkpen, K, (eds.) Human-Computer Interaction. Interact 2021. 226–283. Springer: Cham, Switzerland Julie Butler and Margaret L. Kern. 2016.The perma-profiler: A brief multidimensional measure of flourishing, International Journal of Wellbeing. 6, 1–48 (2016). https://doi.org/10.5502/ijw.v6i3.526.
Wagener, Nadine, Niess, Jasmin, Rogers, Yvonne, Schöning, Johannes (2022) 'Mood Worlds: A Virtual Environment for Autonomous Emotional Expression', CHI proceedings, April-May 2022, New Orleans, available at: https://dl.acm.org/doi/10.1145/3491102.3501861.
Weinel, J. (2019) 'Cyberdream VR: Visualizing Rave Music and Vaporwave in Virtual Reality' In: *Proceedings of the 14th International Audio Mostly Conference: A Journey in Sound*; September 2019; Nottingham.