Enhancing Art Gallery Visitors' Experience through Audio Augmented Reality Technology

Abhraneil Dam[1], Yeaji Lee[1], Arsh Siddiqui[1], Wallace Santos Lages[2], Myounghoon Jeon[1]
[1]Virginia Polytechnic Institute & State University
[2]Northeastern University

Audio Augmented Reality (AAR) applications are gaining traction, especially for entertainment purposes. To that extent, the current study explored its use and effectiveness in enhancing art gallery visitors' experiences. Four paintings were selected and sonified using the Jython algorithm to produce computer generated music (Basic AAR); the audio was then further enhanced with traditional music by a musician (Enhanced AAR). Twenty-six participants experienced each painting in Basic, Enhanced, and No AAR condition. Results show that AAR cues had a significant effect on participants' subjective feedback towards the paintings. Sentiment Analysis shows that participants mentioned significantly more positive words from Enhanced AAR than the others. Enhanced AAR also made participants express a sense of immersion, whereas Basic AAR made them concentrate more on forlorn aspects of the paintings. Findings from this study suggest ways to improve and customize AAR cues for different painting styles, and indicate the need for multi-modal augmentations.

INTRODUCTION

Visual arts or paintings have been a form of expression and communication, with the first works of art dating back to over 45,000 years ago (Brumm et al., 2021). Over the years of evolution, visual art has gone through iterations of different styles of painting–realism, photorealism, expressionism, impressionism, abstract, and surrealism (PictoClub, 2022). Even though paintings have been a substantial element in society, not every visitor is able to fully grasp the meaning behind a painting; in other words, the visual medium might be insufficient to allow all types of visitors to understand and enjoy works of art. Providing additional information about the artwork, through either the same or an alternate modality, could be a way to augment visitors' experiences so they may draw more meaningful conclusions or interpret it more diversely.

Originally, the reality-virtuality continuum was first defined as a spectrum, with reality on the right most end, and a completely virtual reality (VR) existed on the left most end (Milgram & Kishino, Fumio, 1994). On this spectrum, and closer to its right end, is the concept of augmented reality (AR). AR, generally, has been defined as the registered combination of virtual and real-world elements while allowing for real-time interaction (Azuma, 1997; Milgram et al., 1995). Skarbez et al. (2021) analyzed Milgram's reality-virtuality continuum to better describe the dimensions of AR. These were mainly considered in the visual modality, and consisted of Extent of World Knowledge, Immersion, and Coherence. Extent of World Knowledge meant that the system should allow the user to maintain above average situation awareness of their environment, Immersion was a measure of how well the user felt mediated in the augmented version of reality, and Coherence was the measure of how believable the augmentations were. The dimensions for audio AR are similar yet unique from the dimensions of visual AR–Customization, Immersion, and User Context. Customization is the modification of auditory cues for the intended user, Immersion is the extent to which users feel engaged in an augmented reality, and User Context is the extent to which the auditory cues are helpful for the user's primary task (Dam et al., 2022). Auditory cues, that fit these requirements, are a way to augment users' experiences in a visual art gallery using an alternate modality.

Related Works

Museum tours are considered a staple of AAR applications with the expectations that the technology will improve the social aspects of the tour; Bederson (1995) attempted this by developing a system that recognized the user's position relative to the artifact and played descriptive audio through a headset. However, multiple users had to move in sync for the audio to begin and stop playing at the same time for each user. This concept was taken a step further by Vazquez-Alvarez et al. (2016) in which one layer of sonification was audible to users to attract their attention towards respective artifacts, and a second interactive layer of sound started to play as users would get closer to them. Such applications of audio to enhance users' experiences fit the three dimensions of AAR. The sounds are generally Customized for the purposes of a museum visit, the multi-layered sounds improve Immersion, and the Context for them is to allow users to engage with the artifacts more meaningfully during their visit. Vazquez-Alvarez et al. (2012) also tested a similar concept but in an outdoor explorative environment and required active interactions by users to trigger audio cues. AAR applications including passive interactions have also been investigated, where audio cues were made available to users based on strategically placed triggers in the environment (Mynatt et al., 1997). An individual dimension of AAR, such as Immersion was closely correlated to enjoyment (r=0.65) and allowed users to experience both 'parallel and contrasting' worlds (Reid et al., 2005). In the study by Reid et al. (2005), users experienced a variety of sounds through headphones while they moved around inside a square space 150m wide, and experienced an auditory drama, *Riot! 1831*.

With respect to visual artwork, Chang et al. (2014) used a tablet screen to provide additional visual information about paintings; compared to traditional audio commentary, the visual AR-based mobile system led to better learning effectiveness. Another study, employing visual AR cues to provide text-based information over the artwork, measured global learning outcomes; participants wore Google Glasses that guided them through a series of paintings in an art gallery (tom Dieck et al., 2018). The visual AR cues led to participants learning and retaining more factual information about the paintings, compared to a control group that used a paper-based medium to deliver the same information.

Compared to previous studies, the current study is novel in several ways. First, the use of an alternate modality such as audio, to enhance visitors' experiences in an art gallery has not been studied widely in the past. Using an alternate pathway to provide information reduces visual clutter for the visitor, so they are less distracted from the visual medium of the artwork. Secondly, the auditory information being provided in this study does not relate to the factual information about the painting, but instead, is the sonified version of the paintings. This holds the advantage of allowing visitors to better interpret, feel, and immerse themselves more deeply in the artwork. Finally, by using hardware such as bone conduction headphones, visitors are expected to be able to maintain social connections with other gallery visitors, since the headphones allow for outer ears to remain open. Overall, the current study simulates the experiences of

an art gallery in a more controlled environment while still retaining ecological validity to the highest extent possible. Hence, the current study proposes the following research questions:

RQ1: How does sonification, provided as AAR cues, lead to enhanced experiences when viewing artwork?

RQ2: How does enhanced sonification lead to more immersive experiences when viewing artwork compared to basic sonification?

METHOD

Study Design

The study followed a within-subjects single factor design and was conducted in a room where the 2' x 3' prints of four paintings were displayed on three walls. The independent variable was the AAR cue and comprised of three levels - No AAR, Basic AAR, and Enhanced AAR cues. In total, there were 12 (4x3) trials per participant.

Painting Selections

To exhibit different sentiments and art styles, four paintings were selected. "Guernica" (Fig 1a), a Cubist painting portraying the horrors of the second world war, and enshrined in stark colors of white, grey, and black, "The Return" (Fig 1b), in Romantic style, contained rich deep hues of autumn and sunset promoting tristesse, "The School of Athens" (Fig 1c) in Renaissance style exhibiting themes of individuality and humanity, with soft scholarly colors like gold, brown, and white, and "Woman with a Parasol" (Fig 1d) from the Impressionist style cast in thick brush strokes of calm passive colors like green, blue, and white. Figure 1 shows the paintings, and Figure 2 shows the gallery-like setup of the study.

Figure 1. (a) – top left; (b) – bottom left; (c) – center; (d) – right

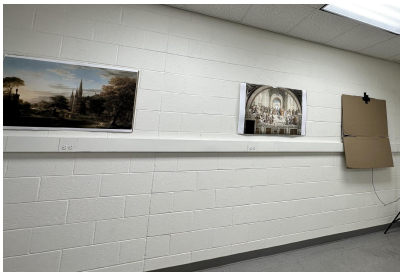

Figure 2. Gallery-like setup

Audio Augmented Reality cues

Basic AAR cues. These were generated in MIDI format from the paintings using a sonification algorithm in Jython Music (Manaris et al., 2016). The major hues and the percentage of each hue in the paintings were identified through a visual parameter analysis at the URL: https://www.geotests.net/couleurs/v2/. These values were then fed to a previously verified sonification algorithm (Nadri et al., 2022) with the Aeolian scale, and with two Pianos and one Percussion as the instruments. Each hue value was mapped to a pitch value in MIDI between 12 and 60 depending on its corresponding percentage; duration of each note was set between 0.75 and 2 seconds, and the volume was set between 30 and 40, on a scale of 0-127. For more in-depth descriptions, please refer to Nadri et al. (2022). This form of AAR allowed participants to view and hear the painting concurrently.

Enhanced AAR cues. Enhancements were made using the software Ableton Live 11 based on the following criteria: if the main colors were considered bright, major scale notes were applied to the MIDI files, and minor scale notes for darker hues. If the painting was 'realist', classical instruments were used, and electronic instruments were used for the non-realist paintings. In paintings with broad brush strokes ("Woman with a Parasol)", reverb was used in the enhancements to provide a dreamlike sensation to accompany the bright sunlight. Similarly, with "Guernica", to provide the effect of not being able to see many colors, electronic sounds were used instead of traditional instruments to provide the effect of not being able to feel the diverse timbre with the sounds. The sounds were also designed to give a sense of space through three-dimensional sound because the main characters are portrayed in a messy manner in the painting so that the participants were able to feel as if they exist in the space of the painting.

All painting and audio files may be viewed here, https://osf.io/ptzvu/?view_only=852429a71027422a96ace8a94b3eeea4.

Participants

Twenty-six (M=11; F=15) participants were recruited with a mean age of 23.4 yrs (*SD*=4.1 yrs). Sample size was calculated with a power of 0.80, significance level of 0.05, and medium effect (η_p^2=0.06). Eligibility criteria required that participants were over 18 years of age with no hearing impairments, and normal or corrected vision. 31% of the participants visited art galleries once a year or once in 3 months, 15% never visited, 12% did so once in six months, 7.5% went once every month, and 3.5% did so every week.

Equipment

In addition to the prints of the paintings, a pair of wireless bone conduction headphone (Aftershokz Bluez 2S) was used in the study. It transmits sound though vibrations in the skull's temporal bone to the inner ear, completely bypassing the outer and middle ear. This allows participants to maintain auditory awareness of their surroundings, and thus, maintain the social aspect of being in an art gallery. This headphone was connected via Bluetooth to a 2021 MacBook Pro 16"; the computer played the Basic AAR cues (MIDI files) using two Steinway Grand Pianos and one SoCal instrument using the GarageBand software, and the Enhanced AAR cues (.wav files) were played using QuickTime player.

Procedure

After signing the consent form, participants filled out a demographic questionnaire. Then, they were allowed to familiarize themselves with the bone conduction headphone. Next, participants were briefed about the experimental protocol as follows: in each trial, participants were instructed to approach one painting, which was covered with a piece of cardboard. One researcher would provide a countdown, at the end of which, another researcher would remove the cardboard and reveal the painting to the participant. As soon as the painting was revealed, the first researcher would also play the corresponding AAR cue for that trial. Participants were asked to swiftly turn away from the painting or make a thumb-up hand gesture to indicate to the researchers that they were done viewing the painting, upon which the first researcher would stop the AAR cue playback. Then, the participant returned to the first researcher to discuss their "thoughts and feelings" about the experience. Their responses were audio recorded and transcribed. This protocol was followed to simulate hardware in which once visitors were within appropriate viewing distance from the painting, the AAR cue would automatically start playing, and when they turned away from the painting, the music would stop playing. Participants were allowed to view the painting

for however long they chose to. Each painting was observed in all three AAR cue types, and the order of presentation was completely randomized for each participant. Sentiment analysis was performed on the audio transcriptions to generate average negative and positive scores for each condition (Medhat et al., 2014). Thematic analysis was also performed on the data using an inductive approach to identify any underlying themes (Nowell et al., 2017).

RESULTS

Sentiment Analysis

The transcribed text was imported into Jupyter in the form of a matrix. The matrix cells contained text for each participant for each painting in each condition. This dataset was analyzed using the SentinemntIntensityAnalyzer library function in Python. After filtering out connecting words in each cell, each remaining word was scored from 0 to 1 for their respective classification of positive, negative, or neutral, with 1 indicating a strong sense for each. Finally, the dataset was separated into two tables, one with the negative scores in each cell, and one with the positive scores in each cell. The positive and negative scores for each condition are summarized in Table 1. An independent samples t-test showed that participants reported significantly greater positive statements (M=0.30; SD=0.03) than negative statements (M=0.13; SD=0.02) after viewing the paintings in the experiment, $t(622) = -12.59$, $p < 0.001$.

Table 1. Average Sentiment Scores by Condition

	Positive Scores		Negative Scores	
	Mean	Std. Dev	Mean	Std. Dev
No AAR	0.31	0.20	0.10	0.17
Basic AAR	0.27	0.18	0.15	0.22
Enhanced AAR	0.33	0.23	0.12	0.17

Next, repeated measures ANOVA test was conducted on the Positive and Negative scores respectively to analyze the effect of the AAR cues on participants' feelings about the paintings. Results showed that there was a significant main effect of AAR cue type on Positive scores $F(2,50) = 3.48$, $p = 0.039$, $\eta_p^2 = 0.122$. A post hoc test with the Bonferroni correction (α=0.05/3 = 0.0167) revealed that there was a significant difference between Basic AAR and Enhanced AAR conditions ($p = 0.011$), but no significant differences between the other conditions. Figure 3 shows mean positive scores for all four paintings; Enhanced AAR had the highest positive scores in all paintings, except for in 'The Return'.

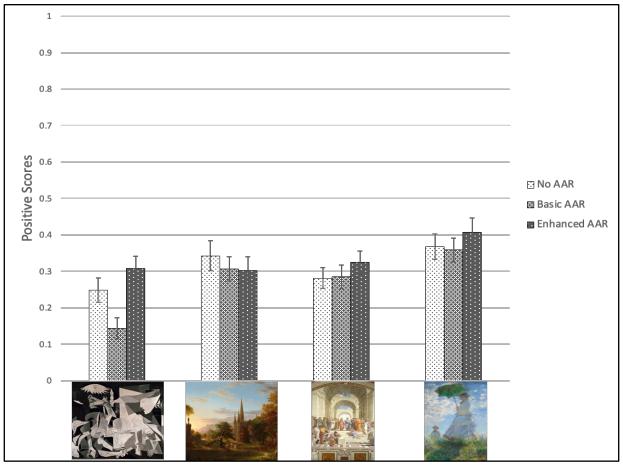

Figure 3. Positive Scores for paintings in all AAR Cue Types (errors bars represent standard error)

Results also showed a significant main effect of AAR cue type on Negative scores $F(2,50) = 3.56$, $p = 0.035$, $\eta_p^2 = 0.126$. Although a post hoc test with Bonferroni correction (α=0.05/3=0.0167) showed results that did not reach the conservative significance level with the adjusted alpha, the tendency was that Basic AAR had higher negative scores than the No AAR ($p = 0.025$), and Enhanced AAR conditions ($p = 0.049$). Figure 4 shows mean negative scores for all four paintings, where Basic AAR had the highest negative scores in all paintings, with the exception of 'The Return'.

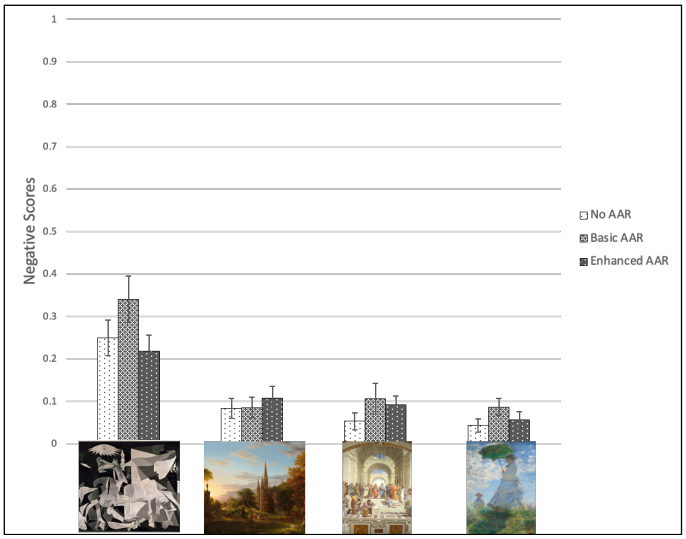

Figure 4. Negative Scores for paintings in all AAR Cue Types (errors bars represent standard error)

Lastly, the feedback from each participant for each painting-audio combination was classified as Supporting or Countering. Statements were classified as 'Supporting' if the words explicitly stated participants' preferences towards the sound-painting combination, or if the words demonstrated a strong or accurate understanding of the painting. Statements were 'Countering' if they explicitly stated disapproval towards the sound-painting combination, or if they displayed a weak or inaccurate interpretation of the painting. In total, there were 242 instances Supporting and Countering statements which were grouped by condition. A single participant could have provided both Supporting and Countering arguments for a painting, in which case their statement was considered in both cases. For Basic AAR cues, there were 49 instances of Supporting and 45 instances of Countering statements. For Enhanced AAR cues, there were 80 instances of Supporting and 45 instances of Countering statements. There were also 23 instances of Supporting statements for the No AAR condition, meaning they preferred to view the painting without music, or provided better description than with AAR cues. Based on these metrics, it can be concluded that Enhanced AAR cues led to 39% greater preference over Basic AAR cues, and a 71% greater preference over No AAR cues. Basic AAR cues led to a 53% greater preference over No AAR cues. Examples of statements for each condition includes:

Supporting: *"I saw the painting inside the painting, and I saw things I didn't' see before, like people are actually fighting…I did not notice that before"* (P1 for 'The School of Athens' in No AAR condition)

Supporting: *"Yeah, that was a good match, music was very sharp, following the sharp edges of the painting, jagged edges, it matches"* (P8 for 'Guernica' in Basic AAR condition)

Supporting: *"The song, I felt, goes well with the painting, like it does create the scene, like a castle in the background…kind of conveys to you the color that used, puts you in that right in that spot right there"* (P25 for 'The Return' in Enhanced AAR condition)

Supporting: *"The s*low, soft, but kind of techno beat… music combined made it seem like kind of relaxing just as if you're waiting for something" (P5 for 'The Woman with a Parasol in Enhanced AAR condition)
Countering: *"Every time there's a beat, I feel like I have to change the frame of my eye, and it's sort of tiring looking at a painting like that.."* (P1 for 'The School of Athens' in Basic AAR condition)
Countering: *"Feels like it's scary, the sound makes you feel like a scary environment"* (P23 for 'The Return' in Enhanced AAR condition).

Emergent Themes

Thematic analysis using an inductive approach also revealed the following themes. These themes support the findings from the Sentiment Analysis, and they are discussed below.

Sense of Darkness with Basic AAR cues. Both Supporting and Countering statements in the basic AAR condition made the recurring point that participants tended to focus on forlorn aspects of the painting or take on a more nihilistic outlook on the painting. There were 24 instances of participants reporting feeling negative emotions and focusing on darker imagery when viewing paintings accompanied by the Basic AAR cues. This was evident through the consistent use of phrases such as, *"serious and scary"*, *"the tone of the notes, just make it seem darker"*, *"noticed the shadows a lot more"*, and *"definitely had a negative view"*. For example, in 'Woman with a Parasol', P14 reported, *"as soon as I started hearing the music, I noticed the storm clouds"*, even though there were only a few strokes of gray color in the sky. In 'The Return', P5 commented, *"since the music is so spread apart, it makes it feels darker than what I seems"*, and *"makes me focus more on the darker parts of the image, like I noticed the guy was injured on the stretcher, and didn't really see that earlier"*; these comments show that despite the painting containing bright colors to depict a sunrise, their focus was on the smaller, more negative details. In "The School of Athens", P11 said, *"you want to look at the painting, but it triggers your attention to like the negative symbols in piece such as the sculptures of people screaming"*. They are referring to a single face that is part of a sculpture in the background of the painting; due to its low saliency and lack of colors, this detail is not easily noticed. However, in the painting 'Guernica', the use of Basic AAR cues caused participants to understand the painting better since the artwork itself is designed to promote negative feelings. Participants reported, *"the music just feels very dark, kind of depressing"* (P5), *"low cords corresponded to the darker tones of the piece"* (P22), *"with this music, I was mainly thinking of darkness and chaos"* (P6), and *"felt even more disturbing"* (P16).

Sense of Immersion with Enhanced AAR cues. When participants viewed the paintings with Enhanced AAR cues, there were 14 instances of participants reporting feeling as though they were a part of the painting. In 'Woman with a Parasol', P25 reported, *"it did give the feeling of wind and the rustling of wings"*, and P22 used *"floating, windy"* to describe the painting. This corroborates with the flowing edges of the woman's dress, and open sky in the painting. When viewing 'The School of Athens', P9 reported, *"the echo that came in the music..made me feel like I was standing in the cathedral"*; P18 said, *"felt like I was front row to this scene"*. This corresponds to the structure in the painting that is depicted with archways. P13 commented, *"felt more real, I could feel it"* when viewing 'The Return', and P18 also remarked, *"I feel like I can just step into the painting"*. This painting showed a vast open space with lots of small details. When viewing 'Guernica', P26 reported, "feeling a shadow come over me", and P21 said, *"lack of sound in areas in the music, made it sound like they are in some sort of dark void"*. This is especially true given the dark nature of the artwork and the use of contrasting black and white colors to portray the horrors of the second world war.

DISCUSSION

The current study tested the effects of AAR cues on visitors' experiences at a simulated art gallery. Results from the sentiment and thematic analysis partially support **RQ1**'s expected findings. Participants reported significantly more negative statements when observing paintings with Basic AAR cues, and expressed an increased focus on forlorn themes within the paintings. This could be explained by the instruments used to play the MIDI files since most Basic AAR cues included low frequency notes from Piano, and bass presence from the SoCal instrument. Despite the random note assignment to each hue based on its percentage value, most of the Basic AAR cues led to participants feeling hopeless, fearful, sorrowful and nihilistic, focusing on similar contents in the paintings, and interpreting the artwork in a similar fashion. It is possible that selecting alternate paintings with greater color diversity and assigning different instruments will lead to fewer negative comments in the future. This was true for all paintings except 'The Return'. Here, the Basic AAR cues were perceived as slightly less negatively than Enhanced AAR cues due to the music consisting of lower frequencies, spread out notes, and a reverb. Forlorn feelings may still be considered effectively enhanced experiences of viewing artwork that purposefully presents imagery in darker hues, or depicts fearful, sorrowful or nihilistic themes, such as in 'Guernica'. Compared to No AAR condition, Basic AAR cues guided participants to focus more on such aspects of the painting, and thus enhanced their interpretations of the work. In the No AAR condition, P14 even commented the painting to *"look kind of humorous"*, which would not be a very accurate interpretation. Hence, we can conclude that generally Basic AAR did lead to more enhanced experiences, with the exception for one painting. Regarding **RQ2**, participants reported significantly greater positive scores and feelings of being immersed in the artwork which supports the expected findings. Musical enhancements made to the Basic AAR cues led to the creation of traditional sounding music, which was more familiar to the participants. It is possible that this familiarity, combined with the absence of low frequency audio in the Enhanced AAR cues, left participants with a more positive opinion of the paintings. This was true for all the paintings except for 'The Return', where the Enhanced AAR cues contained too many low frequency notes and reverb that did not match with the clearly depicted light source and bright colors present in the painting. The sense of immersion that participants experienced also directly corresponds to the Immersion dimension of AAR, and lends further credibility to the same (Dam et al., 2022). Regarding the other two dimensions, the Enhanced AAR cues were better Customized than Basic AAR cues since they were manipulated by a musician to better fit each individual painting; the User Context dimension was also met by the Enhanced AAR cues since the thematic analysis revealed that participants provided a greater percentage of supporting statements about the paintings than they did compared to the Basic AAR (39%) or No AAR cues (71%), indicating that the cues were effective in helping participants to understand the paintings better. Hence, it can be concluded that curated Enhanced AAR cues can lead to more immersive and enhanced experiences compared to Basic and No AAR cues in an art gallery setting.

Findings from the current study align with the user evaluations of the LISTEN system (Zimmermann & Lorenz, 2008) implemented at the Kunstmuseum Bonn exhibiting artworks of painter August Macke. While the auditory content for their system was not related to sonification of the artwork, it provided verbal descriptions and sound collages to visitors through a headset as they approached different paintings; this could be considered an ideal application of the system envisioned in the current study. The LISTEN system was perceived

"positively" by 91% of the 699 visitors, and 65% rated 'Yes' for enriching their viewing experience.

Limitations

Some limitations of this study include the lack of sufficient objective metrics such as time of observation, or eye-tracking metrics. The transformation of subjective feedback into positive and negative words was based on the SentinemntIntensityAnalyzer library; the study should be replicated with additional objective measures to corroborate these findings. The selection of four paintings only also limits the scope of the AAR cues, and it is not possible to extrapolate these results to paintings of other styles or time periods. Finally, since the thematic analysis was conducted and verified between three researchers only, the disadvantages of any thematic analysis might apply to the current study, which is inconsistency and a lack of coherence in the themes (Holloway & Todres, 2003).

Future Work

Future work on this would include adding visual augmentations to the paintings to compare the effects of visual, auditory, and a combination of visual and auditory augmentations. Future studies should also include objective metrics, such as average time spent observing, and eye-tracking metrics (number of gaze points, average gaze duration). Lastly, studies should also consider social aspects of an art gallery setting and recruit multiple/dyadic groups of participants to study interaction patterns, utterances, and discussion topics based on the augmentations.

ACKNOWLEDGEMENT

This research has been supported by the Institute for Creativity Arts and Technology (ICAT) group at Virginia Polytechnic Institute and State University.

REFERENCES

Azuma, R. T. (1997). A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, *6*(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

Bederson, B. B. (1995). Audio augmented reality: A prototype automated tour guide. *Conference Companion on Human Factors in Computing Systems - CHI '95*, 210–211. https://doi.org/10.1145/223355.223526

Brumm, A., Oktaviana, A. A., Burhan, B., Hakim, B., Lebe, R., Zhao, J., Sulistyarto, P. H., Ririmasse, M., Adhityatama, S., Sumantri, I., & Aubert, M. (2021). Oldest cave art found in Sulawesi. *Science Advances*, *7*(3), eabd4648. https://doi.org/10.1126/sciadv.abd4648

Chang, K.-E., Chang, C.-T., Hou, H.-T., Sung, Y.-T., Chao, H.-L., & Lee, C.-M. (2014). Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum. *Computers & Education*, *71*, 185–197. https://doi.org/10.1016/j.compedu.2013.09.022

Dam, A., Siddiqui, A., Leclerq, C., & Jeon, M. (2022). Extracting a Definition and Taxonomy for Audio Augmented Reality (AAR) Using Grounded Theory. *HFES 66th International Annual Meeting Technical Program*, 5. https://doi.org/10.1177/1071181322661434

Holloway, I., & Todres, L. (2003). The status of method: Flexibility, consistency and coherence. Qualitative Research, 3, 345–357. doi:10.1177/1468794103033004

Manaris, B., Stevens, B., & Brown, A. R. (2016). JythonMusic: An environment for teaching algorithmic music composition, dynamic coding and musical performativity. In *Journal of Music, Technology & Education* (Vol. 9, Issue 1, pp. 33–56). Intellect. https://doi.org/10.1386/jmte.9.1.33_1

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

Milgram, P., Haruo Takemura, Akira Utsumi, & Fumio Kishino. (1995). *Augmented reality: A class of displays on the reality-virtuality continuum*. *2351*, 282–292. https://doi.org/10.1117/12.197321

Milgram, P. & Kishino, Fumio. (1994). A TAXONOMY OF MIXED REALITY VISUAL DISPLAYS. *IEICE Transactions on Information Systems*, *Vol E77-D*(No.12), 16.

Mynatt, E. D., Back, M., Want, R., & Frederick, R. (1997). Audio aura: Light-weight audio augmented reality. *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology - UIST '97*, 211–212. https://doi.org/10.1145/263407.264218

Nadri, C., Anaya, C., Yuan, S., & Jeon, M. (2022). From Visual Art to Music: Sonification Can Adapt to Painting Styles and Augment User Experience. *International Journal of Human–Computer Interaction*, 1–13. https://doi.org/10.1080/10447318.2022.2091210

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, *16*(1), 160940691773384. https://doi.org/10.1177/1609406917733847

PictoClub. (2022, September 22). *Types Of Painting Art: Styles, Mediums & Subjects*. https://www.pictoclub.com/types-of-painting-art-styles-mediums-subjects/

Reid, J., Geelhoed, E., Hull, R., Cater, K., & Clayton, B. (2005). Parallel worlds: Immersion in location-based experiences. *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, 1733–1736. https://doi.org/10.1145/1056808.1057009

Skarbez, R., Smith, M., & Whitton, M. C. (2021). Revisiting Milgram and Kishino's Reality-Virtuality Continuum. *Frontiers in Virtual Reality*, *2*, 647997. https://doi.org/10.3389/frvir.2021.647997

tom Dieck, M. C., Jung, T. H., & tom Dieck, D. (2018). Enhancing art gallery visitors' learning experience using wearable augmented reality: Generic learning outcomes perspective. *Current Issues in Tourism*, *21*(17), 2014–2034. https://doi.org/10.1080/13683500.2016.1224818

Vazquez-Alvarez, Y., Aylett, M. P., Brewster, S. A., Jungenfeld, R. V., & Virolainen, A. (2016). Designing Interactions with Multilevel Auditory Displays in Mobile Audio-Augmented Reality. *ACM Transactions on Computer-Human Interaction*, *23*(1), 1–30. https://doi.org/10.1145/2829944

Vazquez-Alvarez, Y., Oakley, I., & Brewster, S. A. (2012). Auditory display design for exploration in mobile audio-augmented reality. *Personal and Ubiquitous Computing*, *16*(8), 987–999. https://doi.org/10.1007/s00779-011-0459-0

Zimmermann, A., & Lorenz, A. (2008). LISTEN: A user-adaptive audio-augmented museum guide. User Modeling and User-Adapted Interaction, 18(5), 389–416. https://doi.org/10.1007/s11257-008-9049-x