When Sound Modulates Vision: VR Applications for Art and Entertainment

Svetlana Bialkova* and Marnix S. van Gisbergen

Academy for Digital Entertainment

NHTV Breda University of Applied Sciences, The Netherlands

ABSTRACT

The interplay between sound and vision is a key determinant of human perception. With the development of Virtual Reality (VR) technologies and their commercial applications, there is emergent need to better understand how audio-visual signals manipulated in virtual environments influence perception and human behaviour. The current study addresses this challenge in simulated VR environments mirroring real life scenarios. In particular, we investigated the parameters that might enhance perception, and thus VR experiences when sound and vision are manipulated.

A VR museum was created mimicking a real art gallery featuring Japanese paintings. Participants were exposed to the gallery via Samsung Gear VR, head mounted display, and could freely "walk" in. To half of the participants newly composed music clips were played, during the VR gallery "visit". The other participants were exposed to the same environment, but no music was played (control condition).

The results showed that music played altered the way people are engaged in, perceive and experience the VR art gallery. Opposite to our expectation, the VR experience was liked more when no music was played. The naturalness and presence were perceived to be relatively high, and did not differ significantly depending on whether music was played or not. Regression modelling further explored the relationship between the parameters hypothesised to influence the VR experiences. The findings are summarised in a theoretical model. The study outcomes could be implemented to successfully develop efficient VR applications for art and entertainment.

Keywords: Sound-vision interplay, art, entertainment.

Index Terms: Augmented and virtual realities, information interfaces and presentation

1 INTRODUCTION

Despite increasing interest in Virtual Reality (VR) and its commercial applicability, it seems the potential of VR applications is not fully explored. From one side, highly instrumented scientific laboratories traditionally perform precision studies, ignoring the customers' needs and demands. From other side, commercial companies, led by market and tech trends, launched a range of consumer grade VR displays, however, often lacking understanding of the fundamental processes underling human perception, and thus the way VR environments are experienced.

Therefore, meeting science and commerce is needed, to rethink the way VR systems are designed and implement [2]. This would not just allow the release of new interactive VR applications, but

---

* email: bialkova.s@nhtv.nl

would offer to consumers the opportunity to experience VR beyond the desktop paradigm, in everyday environments (e.g., home, office, street) as well as to access more futuristic and luxury (VR) environments.

The interplay between sound and vision is such an example where market should meet science and technology, to rethink how to design, develop and implement VR applications. It is a well-known fact that sound and vision are key determinants of human perception. Psychology and psychophysics already extensively explored sound and vision interplay, and its influence on human behaviour, pointing out multifaceted, multisensory integration between these two modalities [16, 17, 20].

Furthermore, neuroscience research has reported a variety of multisensory phenomena in which sensory-specific brain responses and perceptual judgments concerning one sense can be affected in relations with other senses [7]. Thus, the brain seems to work as a multisensory processor in which inputs from different senses complement, modulate, and interact with each other. Note, however, the brain does not just process the input information (e.g., synergetic properties of the physical stimuli) to detect changes in the environment, but optimally resolves ambiguity (between various stimuli and modalities) in the most adaptable fashion [21].

Understanding the mechanism underling the human brain "processor" when it comes to multisensory integrations is still a challenge. Nevertheless, applying fundamental principles of sound and vision interplay, and its influence on human perception and performance inspired the development of games, films, and various applications for entertainment industry. In the creative and entertainment industry it is well known that sound is a powerful tool to communicate meaning, and thus, could reshape cognitive and emotional response, e.g., to a scene and/or a product.

Surprisingly, however, little is known on the role of audio stimuli in VR environments, and in particular, users experience of VR environments mimicking real world scenarios, when interaction between sound and vision may emerge. This knowledge gap calls for further investigation.

The current study addressed the above challenge in a simulated VR environment mirroring a real art gallery. Participants were invited to "visit" the gallery and to freely "walk" in (Samsung Gear VR head mounted display, HMD was used). To better understand how audio-visual signals influence the way virtual environments are perceived, a further manipulation was enrolled. Namely, audio segments were played (via Samsung wireless headphones) to half of the participants while "visiting" the VR gallery. In the control condition, no music was played.

The key drivers of human behaviour while experiencing the VR worlds were further addressed. It turns, however, that on a methodological level there is a gap. Despite the numerous lab studies investigating VR environments, there is no univocal, a sound-proven method measuring the behaviour (change) in response to music played in VR. To appropriately address this

methodological challenge, after performing extensive literature review and some pilots in our lab, we have created a reliable and efficient survey tool (all Cronbach's α's > .70).

The results were unambiguous in showing that sound modulates vision in the VR environment. More precisely, the music played altered the way people are engaged in, perceive and experience the VR art installation.

In the following, we present consecutively the theory behind the study, the method and results. We further provide a discussion on the main findings, from the perspective of existing theories. Finally, we highlight how the outcomes of the current study may be successfully implemented in developing innovative VR applications for creative an entertainment industries.

2 THEORY

2.1 Virtual Reality Environments

Although increasing interest in VR, and despite the capability of computer-simulated 3D environments to create commercial VR applications, it seems their potential has not been fully employed. One reason for limited VR commercial (and everyday use VR) systems is that 3D environments had some technical limitations, including unnatural interface, poor visual quality, lack of interactivity, navigation difficulties [12]. These technical limitations further influence the way VR environments are perceived and experienced.

In this respect, the importance of perceived naturalness of the environment emerged, as a crucial factor driving the virtual worlds experiences [8]. With the advance of VR technology, some of the technical limitations (e.g., realistic high-resolution graphics, and videos) have been resolved, and thus led to increasing the naturalness of the VR worlds.

However, other technical limitations call for further investigation. There is a need to develop better systems that are capable to enhance presence. Presence was defined in terms of subjective experience of being in one place or environment [23], and it could be achieved even when one is physically situated in another environment [11].

As pivotal aspects in enhancing the perceived presence, vividness and interactivity were pointed out [19]. Offering users better displays (e.g., head mounted display), and control devices (e.g., game pads) improved the ability to move freely within the VR environments, as well as the capability to interact with objects. Such interactions further increased engagement.

Engagement is known to be crucial for the way people perceive a movie, play a game, as well as in many other events and experiences designed for/by creative and entertainment industries. In the context of 3D VR environments, engagement reflects the degree of involvement and immersion [11]. A good engagement would provide the feeling of being inside the environment which is experienced by the user. A question therefore arises, how to evoke good engagement and thus to enhance VR experiences?

2.2 Enhancing the Virtual Reality Experience

A plausible way to enhance engagement and thus VR experiences is to enrich the way the environment is presented and thus perceived. Incorporating environmental cues stimulating various modalities (e.g., vision, audition, tactile etc.) would lead to such enhancement. Not surprisingly, improvement of graphical representations in VR was one of the first technical aspects that have been worked on, and thus offering today a good quality visual stimuli to replicate real environments.

Sound was also used in VR, e.g., to create an artificial sense of place [14, 22]. In such a way sound played the role of a vehicle transporting people to somewhere, and thus, creating a sense of being in a place - real, desired, imaginable. While at the infancy of the VR technology, sound quality, quantity and specific sound design patterns for VR and immersion were not available [14], the advance of technology nowadays opens opportunities for new ways to enrich VR sound experiences. Furthermore, the ease, simplicity, as well as costs efficiency of creating virtual environments (in comparisons to real environments and experiences) is worth investing time and research effort. A logical step in our study therefore was to explore the interplay between audio and visual modalities in VR, and the role of music.

2.3 The Role of Music

Aside from the obvious visual (graphics) issues the researchers have been working on to improve, the role of sound, and in particular, music in VR environments seems to do not received their deserved attention. Note however, in creative and entertainment industries, music is a driving force. The notion that music could alter consumers behaviour is not new. Starting early 80's, consumer research dedicated a considerable theoretical interest to the way music influences experiences [5] and choice [13]. Music is capable of evoking complex cognitive and affective responses by consumers. It was reported that music may affect the perception of an environment, e.g., in store when music and lighting were manipulated (changing atmospherics), consumers' experiences varied [3].

Previous work exploring consumers (classical marketing research) and users (usability research) has considered not just how music influences behaviour, but also how different elements in the same environment might interact to produce holistic perceptions of the entire product, environment, experience. Therefore, the next step to advance in the present study is to examine the combined effects of various environmental stimuli, namely sound and vision on people's (users) behaviour and thus the way they experience VR. We hypothesize:

H1: Higher engagement and liking when music is played in the VR environment (in comparison to the control condition - no music).

H2: Naturalness and presence are perceived to be higher when music (than no music) is played.

2.4 Sound and Vision Interplay

Creative and entertainment industries practices showed that the interplay between sound and vision is a driving force. Researchers also provided scientifically proven facts that different sounds influence gaze differently in films/videos [for overview see 15]. The study further showed that the success rate of actually drawing human (visual) attention was determined by the types of sound. In a similar vein, for games, usability testing provided numerous examples when sound modulates vision, and thus determines players engagement [6]. For 3D environments, the interaction between different modalities such as vision and audition was applied to achieve a higher sense of presence [18], as well as to further allow the user to enhance perceived immersion [8]. Moreover, audio and visuals are an important aspect of interactive experiences, which is an extensively used feature within the field of interactive digital games. However, the potential of sound and vision interplay, especially within its applicability in VR is still a very unexplored field at the moment.

Therefore, the current study translate the knowledge on interactive experience, and the above examples into the context of VR. We hypothesize enhanced perceptual response when auditory and visual modalities interplay. The enhanced perceptual response will lead to augmented VR experiences, and thus we assume:

H3: The higher the naturalness and the presence, the higher the engagement will be.

Looking closely at the way the information is processed, further explanation of the interplay between sound and vision was found: Humans react faster to overlapping bimodal audio-visual stimuli than to unimodal (audio or visual) stimuli [16].

2.5 The Human Factor: Multisensory Integration

For the majority of sensory (information) processing, there are sensory specific areas [7]. The visual and auditory systems for example have their separate cortices in the brain, yet also very connected. Some theories further claim that there are certain types of bimodal cells which are responsible for the integration of different sensory inputs, e.g., connecting auditory and visual cues [10].

Note however, the brain does not just process the input information (e.g., synergetic properties of the physical stimuli) to detect changes in the environment. The brain implies refine mechanisms to optimally resolve ambiguity (between various stimuli and modalities) in the most adaptable fashion [21]. Understanding mechanism underling the human brain "processor" when it comes to multisensory integrations is still a challenge. This, however, does not preclude the application of the fundamental principles of sound and vision interplay to VR environments. Moreover, understanding the influence of this interplay on human perception and performance is the corner stone around which we organised the current study.

3 METHOD

To test which are the parameters enhancing perception, and thus VR experiences a VR museum was created, mimicking real art gallery featuring existing Japanese art. In particular, sound and vision interplay was manipulated, while people were "visiting" the VR gallery. Engagement, naturalness and presence were the main variables hypothesised to determine the way R environments are perceived and experienced. We also measured how people liked the VR experience, and their willingness to visit VR and real art galleries (and museums).

3.1 Participants

40 people (16 male and 24 female) took part in the study. The study was conducted in the Netherlands, with EU citizens, mean age 24 years old. All had normal or corrected to normal vision. Only 8 (out of 40) said they have not experienced VR before. 10% of the respondents in the current study reported to do not have knowledge in VR, 35 % reported intermediate knowledge, and 55% reported to have basic knowledge what VR is. Participants were assigned randomly to the experimental conditions, as described below.

3.2 Stimuli and Design

The VR gallery, created by the R&D team of our CRADLE lab featured Japanese art (see Figure 1, for an example of the scene). For each of the paintings exposed in the VR installation, unique music composition was made by professional musicians (our lab industry partners). The music compositions were instrumental only, and reflected the feelings the paintings evoked with the composer.

To test the role of sound and vision interplay on the VR experience, we manipulated the music played. Half of the participants were exposed to music while "visiting" the VR gallery. The other half of the participants "visited" the same VR environment, but no music was played (control condition).

3.3 Procedure and Settings

After a short introduction to the experiment, participants completed a consent form, agreeing to take part in the study. Then they were exposed to the VR environment having the opportunity to experience the art gallery. Participants were informed to have a free "walk" in the gallery as long they want. After having "visited" the VR gallery, they had to complete a questionnaire (see section instrument for details). At the end, participants were debriefed and thanked for the participation.

Perceived engagement, naturalness, presence, liking were the dependent variables measured via a questionnaire, we developed for the purpose of the current study (see section instrument). Willingness to visit real galleries (and museums) was probed, respectively before and after the VR gallery experience.

The VR gallery was created with C# and Unity3D game-engine environments, by our technical team. The developed VR application was run on a Samsung GalaxyS6 smartphone, capable of rendering stereoscopic images and being compatible with the Samsung Gear VR head mounted display (HDM). Audio segments were played via Samsung wireless headphones. A standard gamepad controller was used to navigate through (walk in) the VR art installation.
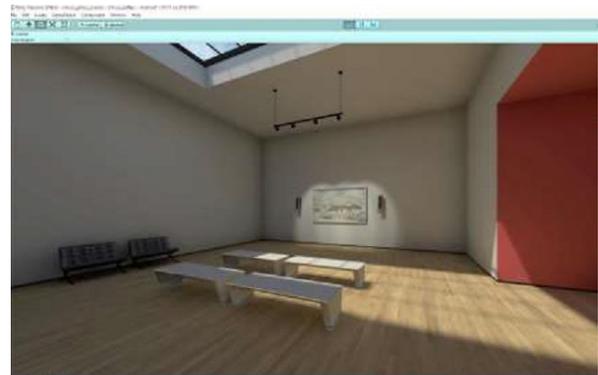
Figure 1: An example of a VR gallery scene.

3.4 Instrument

Taken the novelty of the study, we have developed instrumentarium to measure perceived Engagement, Naturalness and Presence, as well as Liking. The measuring instrument was designed based on extensive literature review and pilot VR studies from our lab. The construct Liking encompassed 3 items (e.g., 'I would have liked the experience to continue'). The construct was developed based on classical marketing literature on liking of products and services (adapted from Aaker, 1996). The Engagement construct encompassed 3 items (e.g., "I felt involved in the displayed environment"). The construct Naturalness encompassed 4 items (e.g., "The displayed environment seemed natural"). The construct Presence encompassed 6 items (e.g., "I felt I was visiting the places in the displayed environment"). Engagement, Naturalness and Presence were designed based on questions adopted from Lessiter, Freeman, Keogh, & Davidoff (2001). All items were measured on a 7-point Likert scale (1 = "strongly disagree", and 7 = "strongly agree").

For the condition where music was played, a 6-items construct captured the behaviour change in response to the music played (e.g., 'The music enriched the VR experience'). The construct is self-developed, based on pilot pre-tests from our lab. All items were measured on a 7-point Likert scale (1 = "strongly disagree", and 7 = "strongly agree").

Few questions captured sociodemographic characteristics (e.g., gender, age, education). We asked about previous experience with VR (no, yes, and if yes which device has been experienced). Knowledge on VR was also addressed (none, basic, intermediate, expert). Frequency of visiting Art galleries, frequency of visiting museums, liking visiting galleries, as well as interest in art and paintings were measured. We also probed the willingness to visit real art galleries (and museums), respectively before and after the VR experience. All items were measured on a 7-point Likert scale (1 = "strongly disagree", and 7 = "strongly agree").

3.5 Analytical procedure

We first performed a reliability check. All scales used demonstrated to be reliable, all Cronbach's α's > .75 (when music was played) and α's > .70 (when music was not played).

T-test compared whether there were significant differences determined by the music played (in comparison to the control condition - no music), respectively for Engagement, Naturalness, Presence, and Liking. The relationship between these 4 variables, hypothesised to determine the way VR environments are perceived and experienced, were further explored with multiple linear regression modelling.

A further test was used to determine whether there was a significant difference in the frequency of visits (of real galleries, museums), as well as to test the willingness of visit real art galleries (and museums), respectively before and after the VR experience.

4 RESULTS

Main effects: VR experience was liked more when no music than music was played ($M = 5.20$ vs. $M = 3.65$), as revealed by a significant T-test outcome, $p < .0001$ (see Figure 2, for an overview of the main effects and Table 1 for sum of the statistics). There was a tendency for deeper engagement when no music than music was played ($M = 4.53$ vs. $M = 3.87$). However, this tendency was not substantiated statistically, $p > .05$. These results are opposite to our expectation for deeper engagement and better liking when music is played, in comparison to the control condition (H1- rejected).

The above outcome is presumably due to the fact that the environment where music was played was perceived as less natural ($M = 4.36$), than the environment where no music was played ($M = 4.73$). As expected, presence was higher when in the VR environment music was played than not ($M = 4.45$ vs. $M = 4.18$). Thus H2 seems to be partially supported. A note of caution is needed here, as the above outcomes are just tendencies. For both, presence and naturalness, the T-tests did not reach significance, p's > .3.
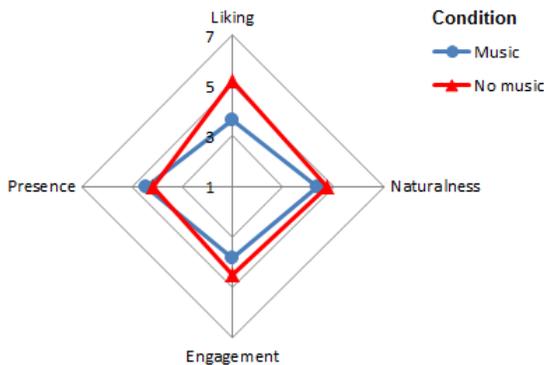

Figure 2: Key parameters influencing VR experiences

Behaviour change in response to music: Participants "visiting" the VR gallery where music was played, reported that music influenced their behaviour ($M = 4.12$). The construct Behaviour change in response to music was highly reliable (Cronbach's α > .8), and it seems to properly capture the way music was experienced, and thus influenced the VR experience. A closer look at the data, showed that the behaviour change in response to music is a function of the perceived naturalness and presence. A regression model encompassing these 2 parameters, explained 29% of the variance in the behaviour change. While perceived naturalness did not have much power ($\beta = .06$), presence contributed significantly to the result observed ($\beta = .52$). Adding engagement improved the explanatory power of the model to 36 %. Respectively, engagement ($\beta = .33$) and presence ($\beta = .34$) contributed significantly to the observed result. The higher the engagement and presence were, the stronger the music modulated behaviour (change). Again, perceived naturalness did not contribute significantly to the explanatory power ($\beta = .05$). We return to this interesting outcome in the general discussion.

Table 1. Summary of the statistics (T-test outcomes)

	Music	NoMusic	t-test
Liking	M=3.65 (SE=.28)	M=5.20 (SE=.24)	t(38) = -4.27, p <.0001
Engagement	M=3.87 (SE=.26)	M=4.53 (SE=.28)	t(38) = -1.76, p = .086
Naturalness	M=4.36 (SE=.25)	M=4.73 (SE=.25)	t(38) = -1.01, p = .321
Presence	M=4.45 (SE=.20)	M=4.18 (SE=.17)	t(38) = 1.01, p = .320

Modelling experiences: The outcomes of the multiple regression modelling clearly suggest that the way participants experienced the VR environment was modulated by the engagement. Moreover, the naturalness and presence influenced engagement. The higher the presence and naturalness were perceived, the higher the engagement was (H3- confirmed).

These modulations varied depending on whether participants experienced the VR environment when music was played or not (see Figure 3, for a summary of the regression modelling, Significant regression coefficients only reported). When music was played, the model encompassing presence ($\beta = .02$) and naturalness ($\beta = .58$), explained 32% of the variance in the engagement, $p < .05$. When music was not played, the same model, with predictors presence ($\beta = .46$) and naturalness ($\beta = .35$), had a better prediction power, explaining 49 % of the variance in the engagement, $p < .005$.

Engagement further influenced liking. The deeper the engagement was, the more the VR experience has been liked. This effect was better pronounced when music was played ($\beta = .83$), in comparison to the control condition ($\beta = .27$).

Behaviour change as a result of the VR experience: To assure that the experiment we have performed had the desired effect, we have asked participants how often they visit galleries in real life, and do the like vising galleries (and museums). Our participants reported that they do not visit that often galleries, and this tendency was relevant no matter to which condition they were assigned ($M = 3.05$ vs. $M = 2.85$, respectively music vs. no music). Participants however showed interest in paintings ($M = 4.65$ vs. $M = 3.95$, respectively music vs. no music).

An interesting trend emerged concerning the willingness to visit real galleries (museums). People were more willing to visit real galleries (museums) after having experienced the VR gallery

featured in the current study. This trend was valid for both conditions, when music was played ($M = 4.5$ vs. $M = 5.6$, before vs. after), and when music was not played ($M = 3.8$ vs. $M = 4.7$, before vs. after).
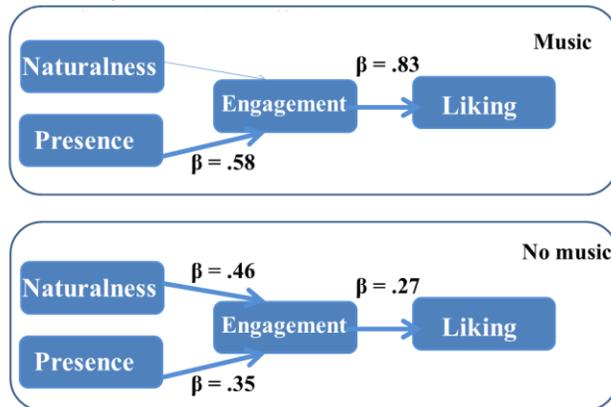

Figure 3: Summary of the regression modeling (Significant regression coefficients only)

5 DISCUSSION

The aim of the current study was to provide a better understanding on how the interplay between audio and visual stimuli in VR environments influence human behaviour (change). In a VR art installation, mimicking real gallery, music was played (or not). We looked at the key parameters determining the way VR environment is perceived and thus experienced. The results are unambiguous in showing that music altered the way people are engaged in, perceive and experience the VR art installation.

5.1 How Sound Modulates Vision

The results confirmed that the interplay between sound and vision is a key determinant of the way people perceive the VR worlds. Opposite to our expectation, however, less engagement and liking were reported when music was played, in comparison to the control condition (H1- rejected). This outcome argues against existing theory for faster performance when bimodal audio-visual stimuli are overlapping, in comparison to when unimodal (audio or visual) stimuli are involved [16, 17]. Nevertheless, a recent study on the role of sound and vision interplay in films/videos, reported that different sounds determine variation in the gaze, and thus perception [15].

Therefore, we had a close look at how the music played (vs control condition) influenced the perceived naturalness and presence. The environment where music was played was perceived as less natural than the environment where no music was played. Presence however was higher when in the VR environment music was played than not, see Figure 2. These two parameters are just shaping tendencies (not substantiated statistically). But if we are allowed to speculate on this outcome, we may say that playing music when exhibiting paintings is not that common, and thus having music in a VR art gallery was somehow unexpected.

The scenario of an art exhibit is a specific scenario without music, unless it is a multimedia style exhibit. Thus, one may also argue that for a particular soundscape (i.e. the sound of an empty room), music can become a distraction when the objective is to simply visualise patterns in a painting. Note, that a muted environment is also unnatural as a soundscape. Thus, it would be interesting for future studies to compare different scenarios,

contexts and soundscapes, and to look at the key parameters emerging hereby, namely naturalness, presence, engagement, liking.

Furthermore, audios in general (being either music or a more abstract soundscape) are very context sensitive. This means that for the same piece of audio different emotions and experiences may be felt, depending on the visual context. Taken that music is capable of evoking complex cognitive and affective responses in consumers [5, 13], it may be even the case that music affected the perception of the VR environment (e.g., gallery atmospherics) differently than the effect it may have in the real life cases. Whether the above are possible scenarios, could be addressed in a follow up study.

5.2 The Role of Music

Participants "visiting" the VR gallery where music was played, reported that music influenced their behaviour. The regression modelling further supported this outcome, demonstrating that engagement and presence contributed significantly to the model explanatory power (36 % of the variance explained). The higher the engagement and presence were, the stronger the music modulated behaviour (change). Again, perceived naturalness did not contribute significantly to the explanatory power.

We have already provided some speculations on the fact that music was not perceived to be natural for the currently simulated environment, namely unexpectedness. Another plausible explanation may be the type of the music played and the emotions the music could evoke. There is a significant body of literature (e.g., psychophysics, psychology, usability) as well as examples from music industry indicating that both, genre and tempo of the music, have differential effects on user mood, perception and behaviour, and that this effect is context dependent. A further study however is needed to closely address the effect of the factors enumerated above on the emotional parameters, when it comes to playing music in VR environments.

In addition, individual users have very different musical preferences, and this is also likely to affect the response. Note the music played hereby resembled Japanese music based on composer interpretation of the paintings and Japanese music in general. This type of music despite being congruent with the content of the paintings, might not be familiar or liked by the visitors who also had little experience with Japanese art (as seen from psychographics questions additionally addressed). As such the outcomes might be different when music played matches the (personalized) interest of the participants. Whether these explanations hold, could be addressed in a follow up study.

This is a worth research, taken that real world music industry is moving away from the commercial practices in physical articles (e.g., records, CDs, DVDs) towards a more interactive form of music utilization. Moreover, the simplicity of creating VR environments, as well as the monetary feasibilities open new avenues researchers and industry should follow.

5.3 Enhancing the Virtual Reality Experience

The way participants experienced the VR environment was modulated by the hypothesised key parameters: engagement, naturalness and presence (see Figure 3 for summary of the regression models). The higher the presence and naturalness were perceived, the deeper the engagement was (H3- confirmed). The good engagement provided the feeling of being inside the experienced environment. This finding coheres with immersion theories associating feeling of belonging to the environment to being "caught in" [4, 11]. While at the infancy of the VR technology, sound design patterns for VR and immersion were not

available [14], the current study demonstrates that sound and vision interplay is a promising tool to enhance perception, and thus to augment VR experience.

Moreover, the newly introduced construct liking paid off. The deeper the engagement was the more the experience was liked. While our participants reported to do not visit that often galleries, they showed a clear interest to visit real life art galleries (and museums) after having experienced the VR gallery featured in the current study. This is a nice outcome encouraging the creation and implementation of VR environments mirroring the real world, and stimulating real museum and art galleries visits.

Taken the advance of VR technology, we further suggest to translate the created VR gallery to another platform allowing to obtain data from the playing experience itself, such as logs, gameplay recordings. Such a new look will empower the methodological potential of sensing users behaviour while experiencing the VR environment. Another pioneering avenue could be the integration of Eye-tracking technology within the VR environment, and thus to directly look at the human-VR interaction, real time. These innovations may not just explore emerging behaviour patterns, but would provide opportunities for creating appropriate VR applications enhancing experiences.

6 CONCLUSION

Emerging VR technology applications are reshaping not just the virtual but also the real worlds. In this respect, rethinking the way VR systems are designed is needed, to better understand how various environmental cues stimulating different modalities (e.g., vision, audition, tactile etc.) influence the way VR environments are experienced. The current study embraced this challenge.

Investigating how sound modulates vision, in simulated VR environments mimicking real life scenarios, we defined key parameters determining how VR experiences could be augmented. The results clearly demonstrate that music played altered the way people are engaged in, perceive and experience VR. The deeper people were engaged in, the better they liked the VR experience. The study also showed that naturalness of the VR environment significantly influenced the perceived presence, engagement, and thus liking. These crucial parameters determining VR experiences should be taken into account to develop creative and entertaining VR applications, crossing boundaries, and thus offering unforgettable VR experiential journey.

REFERENCES

[1] D. A. Aaker. Measuring brand equity across products and markets. *California Management Review,* 38(3):102-120, 1996.

[2] S. Bialkova. Virtual worlds and real market perspectives. *Virtual Revolution 2016* (Veldhoven, The Netehrlands, March 1-3, 2016).

[3] J. Baker, M. Levy, and Dhruv Grewal. An experimental approach to making retail store environmental decisions. *Journal of Retailing, 68*(4): 445–460, 1992.

[4] E. Brown, and P. Cairns. A grounded investigation of game immersion. *In CHI'04 extended abstracts on Human factors in computing systems* pages 1297-1300. ACM 2004.

[5] G. Bruner, and I. Gordon. Music, mood and marketing. *Journal of Marketing, 54*(4): 94–100, 1990.

[6] J.H. Byun, and S. L. Christian. Audial engagement: Effects of game sound on learner engagement in digital game-based learning environments. *Computers in Human Behavior.* 46:129–138. 2015.

[7] J. Driver, and T. Noesselt. Multisensory Interplay Reveals Crossmodal Influences on 'Sensory-Specific' Brain Regions. Neural Responses, and Judgments. *Neuron.* 57:11–23, 2008.

[8] J. Freeman, and J. Lessiter. Here, there and everywhere: The effect of multichannel audio on presence. In *Proceedings of ICAD 2001*, (Espoo, Finland, July 29-August 1, 2001)

[9] J. Lessiter, J. Freeman, E. Keogh, and J. Davidoff. A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence,* 10(3): 282-29, 2001.

[10] J. Lewald, W. H. Ehrenstein, and R. Guski. Spatio-temporal constraints for auditory–visualintegration. *Behavioural brain research* 121(1): 69–79, 2001.

[11] A. McMahan. Immersion, engagement and presence. A method for analyzing 3D video game theory reader, In The Video Game, Theory Reader, Mark J.P. Wolf and Bernard Perron, eds. NY, US: Routledge, Taylor & Francis Group, pages 67- 86, 2003.

[12] E. B. Nash, G.W. Edwards, J. A. Thompson, and W. Barfield. A review of presence and performance in virtual environments. *International Journal of Human-Computer Interaction*, *12*(1): 1-41, 2000.

[13] A. North, D. Hargreaves, and J. McKendrick. In-store music affects product choice, *Nature,* 390: 132, November 1997.

[14] S. Serafin, and G. Serafin. Sound Design to Enhance Presence in Photorealistic Virtual Reality. In *Proceedings of ICAD 2004,* (Sydney, Australia, July 6-9, 2004).

[15] G. Song, D. Pellerin, and J. Granjon. Different types of sounds influence gaze differently in videos. *Journal of Eye Movement Research*, 6(4):1–13, 2013.

[16] C. Spence, and J. Driver. Audiovisual links in exogenous covert spatial orienting. *Perception and Psychophysics*. 59(1): 1–22, 1997.

[17] C. Spence, J. Ranson, and J. Driver Cross-modal selective attention: On the difficulty of ignoring sounds at the locus of visual attention. *Perception and Psychophysics*. 62:410–424, 2000.

[18] R. Storms and M. Zyda, Interactions in perceived quality of auditory-visual display, *Presence.* 9(6): 557– 580, 2000.

[19] J. Steuer. Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, *42*(4): 73-93, 1992.

[20] D. Talsma, D. Senkowski, S. Soto-Faraco, and M. G. Woldorff. The multifaceted interplay between attention and multisensory integration. Trends in Cognitive Science. 14(9): 400-410, 2010.

[21] T. Thesen, J. F. Vibell, G. A. Calvert, and R. A. Osterbauer. Neuroimaging of multisensory processing in vision, audition, touch, and olfaction. Cogn Process. 5: 84–93, 2004.

[22] P. Turner, I.McGregor, S. Turner, and F. Carroll, Evaluating soundscapes as a means to create a sense of presence, In *Proceedings of ICAD 2003*.

[23] B.G. Witmer, and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3): 225-240, 1998.