Music's Interaction with Virtual Environments: Emotion and Perception

Daniel Snow

N00190107

Supervisor: Robert Griffin

Word count:4857

Thesis submitted as a requirement for the degree of MSc in Cyberpsychology, Dun Laoghaire Institute of Art, Design and Technology, 2020/2021

Thesis Declaration

This Thesis is entirely my own work, and has not been previously submitted to this or any other third level institution

28/04/2021

Acknowledgments

I would like to thank my supervisor Robert Griffin and Critical Research lecturer Dr Liam Challenor for their input in developing the study. Ross Turner and Tamsin Snow for providing the music and video, respectively. Iva Bedzula Prebeg and Nora Gasparik for their assistance in coding data. Jeannie O'Brien, Sonny Snow, Frankie Snow, Wayne Cuddy and Brian Snow for their support.

Contents

Abstract

The emotional and perceptual impact of music on visuals has been widely researched. This study explored the opposite relationship and was designed to determine if different visual conditions would influence the perception of and emotional response to a piece of music using a music-specific emotional scale(GEMS 25). Furthermore, the study investigated if different combinations of visuals and music would be perceived as congruent when there was no synchronisation attempt. Participants listened to a single piece of music whilst viewing one of two video conditions that varied in speed and repetitiveness; there was also an audio-only condition. The authors found that visual stimuli impact emotional responses to music but did not influence perception. Additionally, the emotional influence was only significant within one category of the music emotion scale. Video conditions were perceived as congruent despite there being no intentional synchronisation. The experiment looked at complex interactions, and therefore several factors require further investigation. Still, it suggests that digitally manipulated visuals could be an important consideration when investigating the emotional response to music in several settings such as games, music videos and live performances.

The universality of music in human societies is an indicator of its capacity to elicit emotions (Blood & Zatorre, 2001). This potential emotional impact sees music used in diverse areas such as film, games, therapy and advertising (Eerola & Vuoskoski, 2012). The levels at which people are rewarded by music listening varies significantly, from intense pleasure in some to approximately 3-5 % of people with otherwise normal auditory perception, entirely insensitive to it (Martinez-Molina et al., 2019). Music psychology is mainly concerned with what we hear; yet, we seldom consume music isolated from a visual stimulus (Hallam et al., 2016). There is a lack of research examining the influence of visual stimuli on music. The current study examines this relationship, considering emotion and perception.

Music and emotion

The exploration of emotional responses to music is a multi-disciplinary pursuit (Sloboda & Juslin, 2001). From a neuroscience perspective, research has demonstrated that listening to music has correlated physiological responses akin to those experienced during primary rewards(for example, eating). These responses are demonstrated by an increase in dopamine levels and activity in areas of the brain associated with reward behaviours (Blood & Zatorre, 2001). These physiological measures have not answered whether the music causes an increase in dopamine directly or if the emotions induced by music, in turn, result in increased levels (Ferreri et al., 2019). The brain’s reaction to music in the same regions associated with reproduction and survival has led Blood and Zatorre (2001) to hypothesise that evolutionary-based responses may be adapting to allocate meaning to complex sensory inputs, thus allowing for increased pleasure. And while not necessary to survive, music’s ability to induce pleasure can be seen as biologically beneficial to physical and mental health.

In terms of psychology, the definition of emotion and how it is induced by music is not clear cut. It includes both what is felt and expressed (Tan et al., 2017). Moreover, emotions can be perceived through music, with emotional characteristics being communicated to the listener but not directly experienced (Barthet et al., 2012). This complex set of interactions leads to the question of how best to gauge emotional responses to music.

Discrete (or categorical) and dimensional emotion models

The two most significant theoretical models used in research concerning music and emotion are the discrete (or categorical) and dimensional emotion models (Eerola & Vuoskoski, 2011).

The dimensional model argues that the fundamental elements of emotion can be seen as valence and arousal with positive and negative and high to low bipolar dimensions. Arousal accounting for intensity and valence for pleasantness (Harmon-Jones et al., 2017).

Categorical or discrete emotion models posit that there are basic emotions such as sadness, fear, happiness and joy; these are determined in relation to evolutionary drivers (Scherer K.R, 2000).

This evolutionary perspective may call in to question the usefulness of applying a basic emotion model to gauge musically induced emotion. Conversely, as Blood and Zatorre (2001) demonstrated, music activates the same brain regions as primary drivers such as sex and food. So while not directly necessary for survival may still induce similar emotional responses.

Measuring music and emotion

It has been shown that there are emotions that music has difficulty expressing. These emotions tend to be complex and include envy, admiration and shyness. These emotions relate to an object. For example, you can experience anxiety without anchoring it to an object or cause, but this is not possible with an emotion such as envy (Collier, G.l, 2007). These specific characteristics have led several researchers to seek music-centred emotion measures (e.g. Hevner, 1936; Weld, 1912;). Wedin (1972) found three-dimensional scales that were frequently reported while examining emotion expressed in music. Describing them as ‘perceptual- emotional qualities', they were Tension/Energy, Solemnity/Triviality and Gaiety/ Gloom. However, these studies were concerned with the emotion music expressed or conveyed to the listener. In pursuit of a more satisfactory way of measuring music-induced emotion, Zentner et al. (2008) examined the following questions.

Which emotive states are most (and least) frequently induced by music?
How can we adequately classify and measure them?

Two studies were carried out to gather music-specific emotion words and examine the prevalence of experienced and perceived emotions. In the first study, 515 descriptive words were compiled, the sources of these words were 1.) affect terms from five different languages 2.) Terms taken from Clore et al. (1987) Affective lexicon 3.) Words extracted from the literature concerning music and emotion.
146 terms were kept and carried on to the second study. Participants received a booklet containing the 146 terms and a four-point Likert scale (1=never-4 = frequently) and were then asked to rate the terms in relation to their preferred genre. The ratings were completed considering felt emotions and perceived emotions. Nine factors appeared across the felt and perceived conditions and genres. These factors have subcategories of adjectives that ultimately make up the Geneva Music Emotion Scale (GEMS)(Zentner et al., 2008).
The authors of the above study speculate that these emotions derived from musical experiences might have broader implications for how we experience emotions through aesthetic judgment and evaluation(Zentner et al., 2008).

Aesthetics

The arousal of emotion by artistic works has traditionally been categorised in two ways, these are 'everyday emotions' and 'aesthetic emotions' whilst there are disagreements about which is more prominent, a consensus should be possible that both occur to some degree (Juslin, P.N, 2013). An aesthetic experience or object can be differentiated from an everyday object or experience. It can be subdivided into three categories—'Orientational', where the observer is engaged and attentive to the aesthetic source. ‘Cognitive’ where evaluation occurs, symbolism and meaning are created. And thirdly, ‘Affective’ referring to the emotional experience (Markovic .S. 2012).

Virtual environments

It could be suggested that these elements are desirable in terms of increasing presence in virtual environments. Presence can be defined as the sense of being in another place to the one in which you are physically located (Mutterlein, J. 2018).
A specifically developed virtual environment can be viewed as an ‘artificial information environment’(AIE). An environment can be defined as the sum of surroundings and circumstance. How it is navigated is through perception. Moreover, how an AIE is designed, directs and enables how information is interacted with (Wollner, C. 2018). This potential to coordinate information separates an AEI from natural and built environments.
We can occupy virtual environments through various mediums such as Film, Television and computer games (Lehtonen et al., 2005). These usually have audio and visual elements. Digital technologies allow for manipulation of these factors that may, in turn, increase the intensity of an aesthetic experience and therefore potentially, increase presence.
How sound effects presence requires further research (Lum et al., 2018).

Cross-Modal Correspondence

Cross-modal correspondence(CMC) describes a cognitive process in which associations between sensory stimuli are made. An established example is that between pitch (auditory) and height(visual) (Klapetek et al.,2012). Studies have shown CMC occur between pitch and elevation, lightness and movement direction (Spence, 2020). Brunel et al. (2015) studied the cross-modal relationship between lightness (visual) and audio frequencies. Participants were given audio-visual stimuli and tasked with classifying them by lightness of the shape and audio frequency. Light colours and high frequencies, and, dark colours and low frequencies were considered congruent with opposite pairings considered incongruent. The study backed up the previous identification of this CMC (Marks, 1987 ;Klapetek et al., 2012).

A large amount of CMC research has investigated simple elements of auditory stimuli such as timbre, pitch, tempo and mode. The current study uses a piece of music which is a more complex stimulus.

Complex audio visual interactions

As audio technologies were developed and formats such as radio and LP's became consumer products, visual and auditory stimulus separated. Before audio recordings, music would have most likely been accompanied by a physical performance (Thompson et al., 2005).

Studies have shown that the presence of visual and auditory stimuli can have a more significant emotional effect than either condition on their own (Chapados & Levitin, 2008; Vines et al., 2006). However, these studies are concerned with a musical performance in which the visual element is a direct representation of the music. Studies exploring dance (Christensen et al., 2014) have demonstrated similar results. Researchers have tried to establish specific cross-modal patterns between dance and music (Naveda & Leman, 2009; Caramieaux et al., 2009) where physical movements correspond to musical elements.

Music has been found to influence our perception of the spaces we occupy. Yamasaki et al. (2015) created a study that asked people to rate their impression of different environments while listening to music; these were a quiet urban area, a suburban train journey, a busy city environment and a peaceful setting in nature. The results indicate that the music participants were listening to, in the various locations, influenced their description of the environment.

The relationship between music and film (specifically emotion) has been studied extensively, (see Herget A.K., 2019). Music can have the intention to convey emotion, direct attention and provide memory cues (Cohen A.J., 2001). These studies are mostly concerned with music's influence on film. The opposite CMC of visual stimuli's influence on music is a relatively neglected area of research.

Alignment of Accent Structures

A potentially useful framework for examining music and virtual environments' interaction is the use of 'Alignment of Accent Structures'. The theory suggests that music and images do not have to be perfectly synchronised in order for the combined stimulus (Audio and visual) to be perceived as congruent (Lipscomb S.D., 2005). This model of Film Music perception posits that two processes occur when hearing music and viewing film. These are 'An association judgment' and ' A mapping of accent structures'. The association judgment utilises previous events to gauge the appropriateness of the music. The mapping of accent structures pairs point of emphasis from the two modalities (Lipscomb & Kendall, 1994).

Although the model deals specifically with films and their soundtracks, Its applicability to virtual environments such as those seen in gaming and virtual reality is relevant as both these and traditional cinema represent complex visual and auditory conditions. Furthermore, as the complexity of audio and visual inputs increases so does the perception of congruence between them, regardless of levels of synchronisation (Tan et al., 2013). The focus of these studies is music's impact on the film, the reverse relationship is left unexplored.

Visual stimuli's impact on musical perception

Visual and auditory stimuli have been shown to inform emotional affect and cognition. Larger Increases and decreases in emotional tension have been demonstrated in audio-visual versus audio only conditions (Vines et al., 2006). Boltz et al., (2009) conducted a study exploring how visuals impact music perception. Participants were given five pieces of music, visuals accompanied four of the pieces with one presented as an audio only condition. After each piece participants completed various ratings relating to perception of the music. The study showed that the visual stimulus affected musical perception in several ways. Visuals being present had a perceptual effect, rhythms were determined to be busier, melodies faster and louder compared to the audio only presentation.

Conclusion

A review of the current literature shows a scarcity of research concerning visual stimulus impact on music. Studies concern musical performance (Chapados & Levitin, 2008; Vines et al., 2006; Vines et al., 2011) or education (Geringer et al., 1996). In the experiments by Boltz et al. (2009) and Vines et al. (2006) multiple musical excerpts are used. There is a gap in the literature for the analysis of cross modal relationships where visual inputs are independent variables and music is a constant variable. The current study hypothesises that;

H1: The presence of visuals will alter the emotional response to a piece of music.
H2: The presence of visuals will alter the perception of a piece of music.
H3: Unsynchronised videos and music will be perceived as congruent

Method

Participants

Participants were recruited through social media (Facebook, Twitter, Instagram). A total of 135 people took part. The majority had some musical experience (25% advanced, 31% intermediate, 20% beginner). 44% completed the study on a mobile phone (n=60), 39% on a laptop (n= 52) 15% used a monitor (n=20) and 2% used a tablet (n=1.5%). Participants from thirteen countries completed the study ranging in age from 18 to 73 (M = 36, SD=13.92).

Design

Participants were exposed to a piece of music. Following this, respondents answered a questionnaire relating to felt emotions, four Likert like scales measured perception of the music. Two open-ended questions asked participants to describe their experience (See Appendix A for scales and questionnaires).

The study used an experimental between-subject design with participants assigned to only one condition. There were three conditions. These represented the three levels of the independent variable: Visual stimulus (No visuals, audio-only. Smooth and slow video. Video from the previous condition sped up and repeated). Two dependent variables were measured ;emotion and music perception.

Materials

The music used in the experiment was ‘Individuation’ by ‘The Din’. The piece was selected for its medium-slow tempo (61 bpm), repetitive harmonic structure, major modality and relative obscurity (<1000 Spotify plays). These factors have been identified as encouraging a neutral affect (Gabrielson & Lindstrom, 2001). The music is 3:24 in duration and contains no lyrics.

The visual stimulus was a CGI video animation from an art piece entitled 'Showroom' (Snow, 2017). The video does not include any characters such as humans or animals. The location is a non-specific medical facility. Music and narration is included in the original video; however, it was exported with the audio muted. A pre-test of five independent participants rated the video without audio as neutral in terms of affect.

In the second video condition, the animation's speed was increased to 400% and repeated four times. Every second repeat was reversed. The music was combined with the two videos. The audio and video length was matched; however, no other attempts to synchronise the material were performed. All stimuli were the same length (3:24), with the music unaltered (See Appendix B for links to videos).

The Geneva Emotional Music Scale (GEMS 25, adapted from GEMS45) was used to measure emotional responses to the music. The scale was developed to gauge emotions evoked explicitly by music; the authors argue that it is a more appropriate method than the often used two-dimensional model (Zentner et al., 2008). The scale has nine first-order factors (Sadness, tension, Joyful Activation, Power, Peacefulness, Nostalgia, Tenderness, Transcendence, Wonder) and three second-order factors (Sublimity, Vitality, Unease)(Zentner, 2008). In the current study, the Cronbach Alpha Coefficient was .914 . Four 7 point Likert scales measuring perceptual factors were used (Affect, Activity, Tempo and rhythm).

Apparatus

The audio was downloaded as an MP3, and the video as an MP4 and imported into iMovie. A fade was applied to the end of the video so that it aligned with the music. Three videos were uploaded to YouTube in standard definition (Audio only video appears as a black screen. Video condition two, the original video. Video condition three, video at 400% speed with repetition).

Three separate Microsoft Forms links were created representing the three conditions. The three videos were embedded, so participants did not have to follow an external link. All forms were the same except for the embedded video and the exclusion of the open-ended questions in the blank screen condition. The website Allocate.monster was used to create a single link that diverted people to one of the three conditions. Participants used their own screens and headphones. The human ear can hear a frequency range of 31Hz to 17.6 kHz (Heffner & Heffner, 2007). Relatively inexpensive consumer-grade headphones achieve this range. For example, Apple EarPods have a range of 5Hz – 21,000 Hz (Medwetsski et al., 2020).

Procedure

Participants were asked to follow the Allocate Monster web link and were assigned to one of the three conditions; they were informed that the process would take approximately 10 minutes and headphones were necessary to complete the study. The screen that was used for the experiment was identified. The options were mobile, laptop, monitor and tablet. The participants were instructed to view the embedded video in full-screen mode using headphones at a loud but comfortable volume. After exposure to the video, they completed the GEMS 25. The scale contains 25 adjectives (e.g. Calm, Dreamy, Agitated) rated in terms of the level of felt emotion (not the emotion the music expresses) on a Likert scale (1=not at all, 5=very much). After completion of the GEMS 25, four 7 point Likert scales were answered. Affect (1= very negative, 7= very positive). Activity (1= Not active, 7= very active). Tempo (1 = very slow, 7= very fast) and Rhythm (1= very irregular, 7 = very Regular). Then two open-ended questions were asked ‘Do you feel the music and visuals were matched well? Why do you agree or disagree? And ‘Can you describe your experience of the video considering music and imagery?’ The order of scales and questions was consistent among the three conditions to mitigate carry-over effects. Demographic questions were asked at the end of the study to reduce participant fatigue when viewing the video.

Ethics

Ethical approval was applied for and granted by the IADT Department of Technology and Psychology Ethics Committee through the ‘Amber Route’ (G. Kirwin, personal communication, December 04, 2020). The experiment and surveys satisfied the requirement to be ‘highly unlikely to cause any harm or distress to participants’. The study required all participants to be over 18 years of age. A unique ID code was created by each respondent allowing for removal of their data if requested (See Appendix C for ethical approval and participant information)

Results

Emotion was scored in terms of the nine first factor orders of the GEMS 25 (Zentner, 2008). The range of scores for, Transcendence, Power, Tenderness, Nostalgia and Joyful activation was 3 to 15 and for Sadness and Tension, it was 2 to 10. The first-factor emotion scores are summarised in Table 1. Second-factor emotions were also scored and are summarised in Table 2. The possible range of scores for Unease is 3 to 15 ;Vitality 5 to 25 and Sublimity 11 to 55. The sample was not normally distributed (see tables 1 and 2).Nonparametric tests were utilised.

Table 1. First factor emotions

	N	*M*	*SD*	Skewness	Kurtosis
Tension	132	3.55	2.13	1.47	1.21
Sadness	132	3.16	1.45	1.18	.52
Joyful Activation	130	4.86	1.91	1.53	3.04
Peacefulness	132	8.90	3.19	-.04	-.93
Nostalgia	131	7.70	3.30	.52	-.86
Tenderness	131	7.18	2.91	.36	-.68
Power	133	5.10	2.39	1.13	1.04
Transcendence	128	7.40	2.51	.09	-.57
Wonder	131	7.28	2.85	.27	-.77

Table 2. Second Factor emotions

	N	*M*	*SD*	Skewness	Kurtosis
Sublimity	124	29.7	10.1	.30	-.79
Unease	132	6.71	2.86	1.34	1.78
Vitality	130	9.98	3.96	1.37	2.72

Separate Kruskal-Wallis H Tests were carried out between the three levels of the independent variable and the nine first and three second-factor emotional categories. A further Kruskal-Wallis H test was conducted between the three levels of the Independent variable, and the affect Likert scale (1= very negative, 7= very positive).

A significant result was identified between the first-factor emotion of Peacefulness across the three visual conditions (No video, n = 48 : Original video, n = 38: Repeated video, n=49: p = .005). Post hoc tests show significant results between the no video condition (Md = 9.5) and the repeated video (Md = 7,p = .014). The original video (Md = 10) and the repeated video (p = .003). However, there was no significant result between the no video and original video conditions (p = .461).

No other significant results were identified between the three visual conditions and first factor emotional categories (Tension p = .437: Sadness p = .431: Joyful activation p = .701; Nostalgia p = .063: Tenderness .129: Power p = .293: Transcendence p = .094 : Wonder p = . 212) or second factor emotional categories (Sublimity p = .111: Vitality p = .463: Unease p = .516)

The Kruskal-Wallis H test between visual stimulus and the affect Likert scale identified a significant result in the repeated video (Md = 4) and original video (Md = 5) comparison (p = .018). No significance was found between the repeated video and original video (p=.065) or the audio only and original video (p = .542).

In order to test the hypothesis concerning visual stimulus impact on music perception (H2), three separate Kruskal-Wallis H tests were carried out between the three experimental conditions and the three 7-point Likert scales (Activity, 1=not active at all, 7 = very active: Tempo, 1= very slow, 7 = very fast: Rhythm, 1= very regular). No significant results were returned (Activity p = .847: Tempo p = .351: Tempo p = .180).

A deductive content analysis was carried out on the open ended question 'Do you feel the music and video were matched well?' answered by the participants in the original video (N=38) and repeated video conditions (N=49). Answers were categorised as either agree, disagree, or neither agree or disagree. The responses 'yes', 'agree', 'well matched' 'Matched well' were coded as agree. The responses 'Disagree' , 'no', 'mismatched' 'not well matched' 'didn't match' were coded as disagree. Responses that contained clarifiers such as ‘somewhat’ ,‘initially’, ‘somewhat matched’, ‘one way or the other’ ‘moderately’ were coded as neither agree or disagree. Two responses were removed from the repeated video condition as the responses did not address the question.

Two independent coders rated the same responses (n = 86), using the same three priori categories of ‘Agree’, ‘disagree’ and ‘Neither agree or disagree ‘there was a mean agreement of 90% across the two conditions demonstrating reliability of the scoring (see Appendix E).

In the original video condition 31 (82%) responses were categorised as ‘agree’. 2 responses (5%) were categorised as disagree and the remainder(n = 5 ;13%) were in the neither agree or disagree category.

Analysis of the answers in the repeated video condition show Agree accounted for 60% of responses (n = 28), Disagree 19% (n = 9), and neither agree or disagree 21% (n = 10).

Discussion

The dramatic increase in multimedia information via new technologies will see quantities of data grow, conversely new models will be required to investigate this information and how we interact with it (Sebe, 2010).

Emotion is often the factor that separates human to human and human computer interaction. Human to human communication requires an understanding of the combination of audio and visual stimuli (Jaimes & Sebe, 2007). By developing paradigms of CMC in virtual contexts there is potential to make more meaningful human computer interactions.

The findings of this research show that visual stimuli can influence an emotional response to music. There are several factors to consider. A significant result was found pertaining to one of the nine first factor orders of the GEMS 25 (Zentner, 2008). Peacefulness was a total score for the adjectives; serene, calmed and soothed. Moreover, no significant result was identified between the no video and the original video condition, only between the original video, which showed a higher score for peacefulness than the repeated video, and no video condition, which had a higher score for peacefulness than the repeated video conditions. This implies that the music was considered congruent with the adjectives calm, serene and soothed and the repeated video influenced participant's responses away from these emotions.

Concerning the repeated video condition's influence, the difference in scores was greater between the original video and repeated video than the audio-only and repeated groups, suggesting the original video did have some effect. Although there was only one significant result in terms of emotional categorisation, it is important to highlight that intensity of emotion can vary even with a considered selection of visual stimuli, with some emotions hard to evoke (Gross & Levenson, 1995). Previous studies have shown the induction of peace can be particularly strong (Kreutz et al., 2008). The variation of intensity could explain why the sample did not have a normal distribution and, in turn, highlights that subtler variations may have been missed due to the necessity of non-parametric tests. A curious finding of previous research is that a piece of music can influence various characters in a film differently; it is hypothesised that varying movement of visuals creates different points of emphasis, the more congruent the music and visuals, the higher the level of attention (Cohen, 2015). The content analysis conducted on the open-ended questions shows a higher level of perception of congruence between the original video and music over the repeated video and music, which also indicates more ambiguity in terms of congruence. This difference in perception of congruence could be explained by the Alignment of Accent Structures model, which posits that how notable moments in music and visuals are paired lead to significant differences in perception of the combined audio and video (Lipscomb, 1995)

Three factors differentiate the two video conditions; these are speed, reversed footage and repetition. The current study suggests that these alterations reduced a sense of peacefulness and increase negative responses. However, the experiment design does not allow for further insight into the factors that had an impact. It is possible to speculate that the speed and rhythmic nature of the repeated video led to an increase in synchronisation errors, reducing the perception of congruence; however, further investigation is required.

A significant result was shown in terms of overall affect (negative to positive); again, this was between the original video (more positive) and repeated video (more negative) conditions; however, not between the no video and repeated video conditions. These results show that visuals are not enough to have a significant impact on their own. The finding that there was no significance in terms of affect between the audio-only and repeated video conditions implies that the original video did have some positive effect combined with an increased negative response to the repeated condition resulting in a significant finding. Previous research has demonstrated an increase in emotional processing when audio and visuals are combined (Baumgartner et al., 2006).

No significant results were returned concerning the three perceptual factors (Tempo, Rhythm, Activity. This finding does not support previous research where visual information regardless of congruence has resulted in the perception of increased tempo, more diverse rhythms and higher levels of activity (Boltz et al., 2009). A key difference in the current study is the use of one piece of music and one video. The above findings concerned multiple short excerpts which have not been replicated with the more prolonged, more complex stimuli used in the current study. A review of the existing literature has found no studies that have explored visuals impact on music using entire pieces. The presented findings show significant results that require further research.

Strength and Limitations

The most significant limitation of the current study is a lack of control of experimental conditions. As the study was conducted online, it was impossible to plan for external distractions such as noise or interruptions. The study relies on participants following guidance but cannot monitor if these are being adhered to; For example, whether full-screen mode was entered for viewing the video.
Device type was collected to determine if different screen sizes had an impact. Analysis of the data did not show significant differences between these groups and emotional or perceptual response. A further limitation was not accounting for musical preference or surroundings. Complex relationships exist between context, audience and music (Song, 2016). The scale used in the experiment is a brief version which the authors note is inferior both qualitatively and psychometrically to the full 45 question scale (Zentner, 2008). The reduced scale was used to mitigate participant fatigue.
There were significant benefits to the design. The concise nature of embedded video and easy accessibility encouraged a larger sample size than may have been otherwise possible. As stated, screen type did not significantly impact findings, and participants' headphones are unlikely to have affected frequency response. The GEMS 25 Cronbach Alpha shows good reliability and internal consistency.

Further research

Further research could utilise a within subjects' design using controlled surroundings (screen, headphones, environment). Also, additional qualitative data should be collected to offer a richer insight and identify themes that emotional scales may not recognise. These approaches would address the above limitations.

Virtual reality could be used for future research as it allows for standardisation of environments and manipulations that are not possible in the real world (Roberts et al., 2019). It has been proposed that emotions can be tracked and manipulated in virtual worlds (Chirico et al., 2016).
It allows for a level of interaction that is impossible with audio and 2D video setups. This interactivity opens up the possibility of haptic feedback that could be used to identify/create specific perceived accent structures—offering further insight into perception and cognition.

Conclusion

The current study offers insight into an area of research that has received minimal attention, presenting significant results that can be built upon to increase the understanding of complex audio and visual interactions. As research has focused on simple stimuli and how they interact, it has neglected the complex interactions that occur and are increasingly part of the technology we use. By examining not only perceptual but emotional impacts, we can better understand both positive and negative aspects of technological developments.

References

Barthet, M., Fazekas, G., & Sandler, M. (2012, June). Music emotion recognition: From content-to context-based models. In *International Symposium on Computer Music Modeling and Retrieval* (pp. 228-252). Springer, Berlin, Heidelberg.

Baumgartner, T., Lutz, K., Schmidt, C. F., & Jäncke, L. (2006). The emotional power of music: how music enhances the feeling of affective pictures. *Brain research*, *1075*(1), 151-164.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the national academy of sciences*, *98*(20), 11818-11823.

Boltz, M. G., Ebendorf, B., & Field, B. (2009). Audiovisual interactions: The impact of visual information on music perception and memory. *Music Perception*, *27*(1), 43-59.

Brunel, L., Carvalho, P. F., & Goldstone, R. L. (2015). It does belong together: cross-modal correspondences influence cross-modal integration during perceptual learning. *Frontiers in psychology*, *6*, 358.

Caramiaux, B., Bevilacqua, F., & Schnell, N. (2009, February). Towards a gesture-sound cross-modal analysis. In *International Gesture Workshop* (pp. 158-170). Springer, Berlin, Heidelberg.

Chapados, C., & Levitin, D. J. (2008). Cross-modal interactions in the experience of musical performances: Physiological correlates. *Cognition*, *108*(3), 639-651.

Chirico, A., Yaden, D. B., Riva, G., & Gaggioli, A. (2016). The potential of virtual reality for the investigation of awe. *Frontiers in psychology*, *7*, 1766.

Christensen, J. F., Gaigg, S. B., Gomila, A., Oke, P., & Calvo-Merino, B. (2014). Enhancing emotional experiences to dance through music: the role of valence and arousal in the cross-modal bias. *Frontiers in human neuroscience*, *8*, 757.

Clore, G., Ortony, A., & Foss, M. (1987). The psychological foundations of the affective lexicon. *Journal of Personality and Social Psychology, 53,* 751–766.

Cohen, A. J. (2015). Congruence-association model and experiments in film music: Toward interdisciplinary collaboration. *Music and the Moving Image*, *8*(2), 5-24.

Cohen, A. J. (2001). Music as a source of emotion in film. *Music and emotion: Theory and research*, 249-272.

Collier, G. L. (2007). Beyond valence and activity in the emotional connotations of music. *Psychology of Music*, *35*(1), 110-131.

Eerola, T., & Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, *39*(1), 18-49.

Eerola, T., & Vuoskoski, J. K. (2012). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Perception: An Interdisciplinary Journal*, *30*(3), 307-340.

Ferreri, L., Mas-Herrero, E., Zatorre, R. J., Ripollés, P., Gomez-Andres, A., Alicart, H., ... & Rodriguez-Fornells, A. (2019). Dopamine modulates the reward experiences elicited by music. *Proceedings of the National Academy of Sciences*, *116*(9), 3793-3798.

Geringer, John M., Jane W. Cassidy, and James L. Byo. "Effects of music with video on responses of nonmusic majors: An exploratory study." *Journal of Research in Music Education* 44, no.3 (1996): 240-251.

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & emotion*, *9*(1), 87-108.

Hallam, S., Cross, I., & Thaut, M. (Eds.). (2016). *Oxford library of psychology. The Oxford handbook of music psychology (2nd ed.).* Oxford University Press

Harmon-Jones, E., Harmon-Jones, C., & Summerell, E. (2017). On the importance of both dimensional and discrete models of emotion. *Behavioral sciences*, *7*(4), 66.

Heffner, H. E., & Heffner, R. S. (2007). Hearing ranges of laboratory animals. *Journal of the American Association for Laboratory Animal Science*, *46*(1), 20-22.

Herget, A. K. (2019). On music's potential to convey meaning in film: a systematic review of empirical evidence. *Psychology of Music*, 0305735619835019.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology, 48,* 248–268.

Jaimes, A., & Sebe, N. (2007). Multimodal human–computer interaction: A survey. *Computer vision and image understanding*, *108*(1-2), 116-134.

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of life reviews*, *10*(3), 235-266.

Klapetek, A., Ngo, M.K. & Spence, C. Does crossmodal correspondence modulate the facilitatory effect of auditory cues on visual search?. *Atten Percept Psychophys* **74,** 1154–1167 (2012). https://doi.org/10.3758/s13414-012-0317-9

Kreutz, G., Ott, U., Teichmann, D., Osawa, P., & Vaitl, D. (2008). Using music to induce emotions: Influences of musical preference and absorption. *Psychology of music*, *36*(1), 101-126.

Lehtonen, M., Page, T., & Thorsteinsson, G. (2005, December). Emotionality Considerations in Virtual Reality and Simulation Based Learning. In *CELDA* (pp. 26-36).

Lipscomb, S. D. (2005). The perception of audio-visual composites: Accent structure alignment of simple stimuli. *Selected reports in Ethnomusicology*, *12*, 37-67.

Lipscomb, S. D. (1995). *Cognition of musical and visual accent structure alignment in film and animation* (Doctoral dissertation, University of California, Los Angeles).

Lipscomb, S. D., & Kendall, R. A. (1994). Perceptual judgement of the relationship between musical and visual components in film. *Psychomusicology: A Journal of Research in Music Cognition*, *13*(1-2), 60.

Lum, H. C., Greatbatch, R., Waldfogle, G., & Benedict, J. (2018, September). How immersion, presence, emotion, & workload differ in virtual reality and traditional game mediums. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 62, No. 1, pp. 1474-1478). Sage CA: Los Angeles, CA: SAGE Publications.

Marković, S. (2012). Components of aesthetic experience: aesthetic fascination, aesthetic appraisal, and aesthetic emotion. *i-Perception*, *3*(1), 1-17.

Marks, L. E. (1987). On cross-modal similarity: auditory–visual interactions in speeded discrimination. *Journal of experimental psychology: Human Perception and Performance*, *13*(3), 384.

Martínez-Molina, N., Mas-Herrero, E., Rodríguez-Fornells, A., Zatorre, R. J., & Marco-Pallarés, J. (2019). White matter microstructure reflects individual differences in music reward sensitivity. *Journal of Neuroscience*, *39*(25), 5018-5027.

Medwetsky, L., Kelly, K., & Bakke, M. (2020). Earphone models for iPhones: surprising results when used with a hearing app. *Accessed September*, *24*.

Mütterlein, J. (2018, January). The three pillars of virtual reality? Investigating the roles of immersion, presence, and interactivity. In *Proceedings of the 51st Hawaii international conference on system sciences*.

Naveda, L., & Leman, M. (2009). A cross-modal heuristic for periodic pattern analysis of samba music and dance. *Journal of New Music Research*, *38*(3), 255-283.

Roberts, A. C., Yeap, Y. W., Seah, H. S., Chan, E., Soh, C.-K., & Christopoulos, G. I. (2019). Assessing the suitability of virtual reality for psychological testing. *Psychological Assessment, 31*(3), 318–328. https://doi.org/10.1037/pas0000663

Sebe, N. (2010). Human-centered computing. In *Handbook of ambient intelligence and smart environments* (pp. 349-370). Springer, Boston, MA.

Scherer, K. R. (2000). Psychological models of emotion. *The neuropsychology of emotion*, *137*(3), 137-162.

Sloboda, J. A., & Juslin, P. N. (2001). Psychological perspectives on music and emotion. *Music and emotion: Theory and research*, 71-104.

Snow, T.(2017) Showroom [Video] https://www.tamsinsnow.com/showroom.html

Song, Y. (2016). *The Role of Emotion and Context in Musical Preference* (Doctoral dissertation, Queen Mary University of London).

Spence, C. (2020). Simple and complex crossmodal correspondences involving audition. *Acoustical Science and Technology*, *41*(1), 6-12.

Tan, S. L., Cohen, A. J., Lipscomb, S. D., & Kendall, R. A. (Eds.). (2013). The psychology of music in multimedia.

Tan, S. L., Pfordresher, P., & Harré, R. (2017). *Psychology of music: From sound to significance*. Routledge.

Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, *2005*(156), 203-227.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*(1), 80-113.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., & Levitin, D. J. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, *118*(2), 157-170.

Wedin, L. (1972). A multidimensional study of perceptual-emotional qualities in music. *Scandinavian journal of psychology*, *13*(1), 241-257.

Weld, H. P. (1912). An experimental study of musical enjoyment. *Amer- ican Journal of Psychology, 23,* 245–308.

Wöllner Clemens. (2018). *Body, sound and space in music and beyond: multimodal explorations*. Abingdon, Oxfordshire: Routledge

Yamasaki, T., Yamada, K., & Laukka, P. (2015). Viewing the world through the prism of music: Effects of music on perceptions of the environment. *Psychology of Music, 43*(1), 61–74. https://doi.org/10.1177/0305735613493954

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion*, *8*(4), 494.

Appendix A

Scales and questionnaires

Figure A-1
GEMS 25

Geneva Emotional Music Scale (GEMS-25)

Instructions
When providing your ratings, please describe how the music you listen to makes you *feel* (e.g., this music makes me *feel* sad). Do not describe the music (e.g., this music is sad) or what the music may be expressive of (e.g. this music expresses sadness). Bear in mind that a piece of music can be sad or can sound sad without making you feel sad. Please rate the intensity with which you felt each of the following feelings on a scale ranging from 1 (*not at all*) to 5 (*very much*).

1	**2**	**3**	**4**	**5**
Not at all	Somewhat	Moderately	Quite a lot	Very Much

1. _____ Moved
2. _____ Fascinated
3. _____ Strong
4. _____ Tender
5. _____ Nostalgic
6. _____ Serene
7. _____ Animated
8. _____ Sad
9. _____ Tense
10. _____ Bouncy
11. _____ Filled with Wonder
12. _____ Sentimental
13. _____ Affectionate
14. _____ Overwhelmed
15. _____ Agitated
16. _____ Feeling of Transcendence
17. _____ Calm
18. _____ Joyful
19. _____ Tearful
20. _____ Soothed
21. _____ Energetic
22. _____ Dreamy
23. _____ Mellowed (Softened-up)
24. _____ Allured
25. _____ Triumphant

Information for Researchers

The GEMS-25 is the brief version of the GEMS-45. It is based on the same adjective items pool and the same data-analytic procedures reported in our original article, in particular in Appendix A (see Zentner et al., 2008). The only difference being that we aimed at fitting the model with a lesser number of items so as to provide a shorter measure of the nine musical emotion factors. The best items, i.e. those having relatively high factor loadings all while maximizing model-fit, were the ones retained for GEMS-25. Still, it needs to be kept in mind that the brief version is both psychometrically and qualitatively inferior to the full 45-item version of the GEMS. When using this scale, please carefully read the "general remarks" in Appendix A of our original article, in which certain properties of the GEMS are noted. Participants may be encouraged to add terms they think are missing from the list in a space reserved for this purpose (e.g., after item 25).

Scoring Instructions for GEMS-25

Items belonging to each musical emotion category are as follows:

Wonder: 1, 11, 24
Transcendence: 2, 14, 16
Power: 3, 21, 25
Tenderness: 4, 13, 23
Nostalgia: 5, 12, 22
Peacefulness: 6, 17, 20
Joyful Activation: 7, 10, 18
Sadness: 8, 19
Tension: 9, 15

Feedback and Findings

We are extremely interested in hearing about your experiences with this instrument and to be informed about relevant findings or publications.

Reference

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, 8, 494-521.

Contact: Please send all enquiries to: marcel.zentner@uibk.ac.at

Figure A-2
Perceptual Likert scales

6

Affect : 1 = Very negative, 7 = Very positive
Activity: 1 = Not active at all, 7 = Very active
Tempo: 1 = Very slow, 7 = Very fast
Rhythm 1 = Very irregular, 7 = Very regular

	1	2	3	4	5	6	7
Affect	○	○	○	○	○	○	○
Activity	○	○	○	○	○	○	○
Tempo(speed of music)	○	○	○	○	○	○	○
Rhythm	○	○	○	○	○	○	○

Figure A-3
Open ended questions

7

Do you feel the music and visuals were matched well? Why do you agree or disagree?

Enter your answer

8

Can you describe your experience of the video considering music and imagery?

Enter your answer

Appendix B

Video links and screenshots

Links to video conditions

Video Original - https://youtu.be/RUwCbAV7kKI

Video Repeated - https://youtu.be/HSVS1z7UPRs

Video No Visuals - https://youtu.be/l3ny9Elk_vw

Figure B-1
Screenshot from video

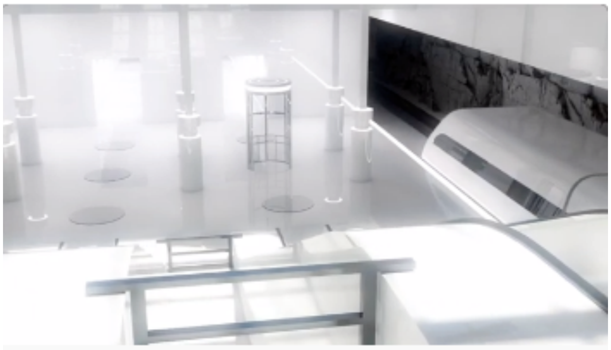
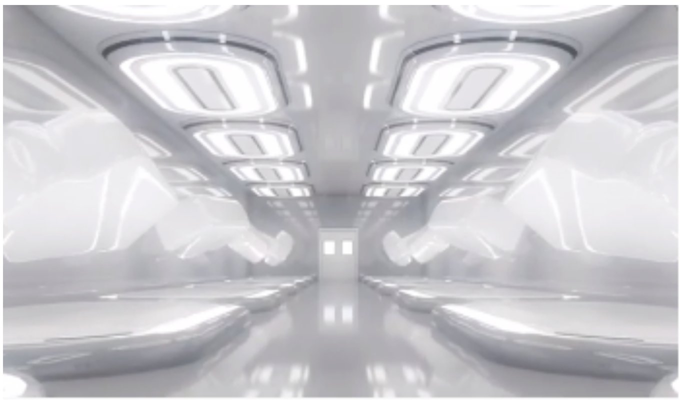


Figure B2
Screenshot from video2



Appendix C

Ethics and Information for Participants

Figure C-1

Ethical approval

Dear Daniel

Thank you for your Amber ethics application for the project entitled "Cross modal correspondences Emotion and perception of music "

I am pleased to inform you that your application has been approved and you can begin data collection when your supervisor advises that it is appropriate to do so.

Best of luck with your research

[redacted]

[redacted] CPsychol

Lecturer in Psychology

Department of Technology and Psychology
IADT
Kill Avenue, Dun Laoghaire, Co. Dublin, Rep. of Ireland, A96 KH79
[redacted]

Figure C-2
Participant information sheet

Purpose of the Research
The research aims to explore the relationship between auditory and visual stimuli in virtual environments.

Invitation
You are being invited to consider taking part in the research study 'Virtual environments influence musical perception'. Daniel Snow is undertaking this project.

Before you decide whether or not you wish to participate, it is important for you to understand why this research is being done and what it will involve. Please take time to read this information carefully and discuss it with friends and relatives if you wish. Ask us if there is anything that is unclear or if you would like more information.

Do I have to take part?
You are free to decide whether you wish to take part or not. If you decide to take part, you will be asked to sign two consent forms; one is for you to keep and the other is for our records. You are free to withdraw from this study at any time and without giving reasons.

If I take part, what do I have to do?
You will complete a short questionnaire that will take no more than two minutes to complete. You will then watch a video accompanied by instrumental music or listen to a piece of music without a video. The video and music are three minutes and thirty seconds long. Following the video, you will complete a questionnaire. The whole experiment will last for approximately 10 minutes.

What are the benefits of taking part?
You will be participating in an experiment that will potentially contribute to knowledge in music perception and how visual and auditory senses interact.

What are the disadvantages and risks of taking part?
There are no inherent risks in taking part. If you wish to end the experiment at any point, you are free to do so.

How will information about me be used?
Your data will be treated with total confidentiality. There is no benefit to the study to associate data with a name. The data will be stored anonymously; a participant number will be assigned at random. Data will be stored on a password-protected hard drive. The results of the research form part of a thesis for an MSc in cyberpsychology at IADT Dublin.

Who has reviewed the study?
This study has been approved by the Department of Technology and Psychology Ethics Committee (DTPEC).

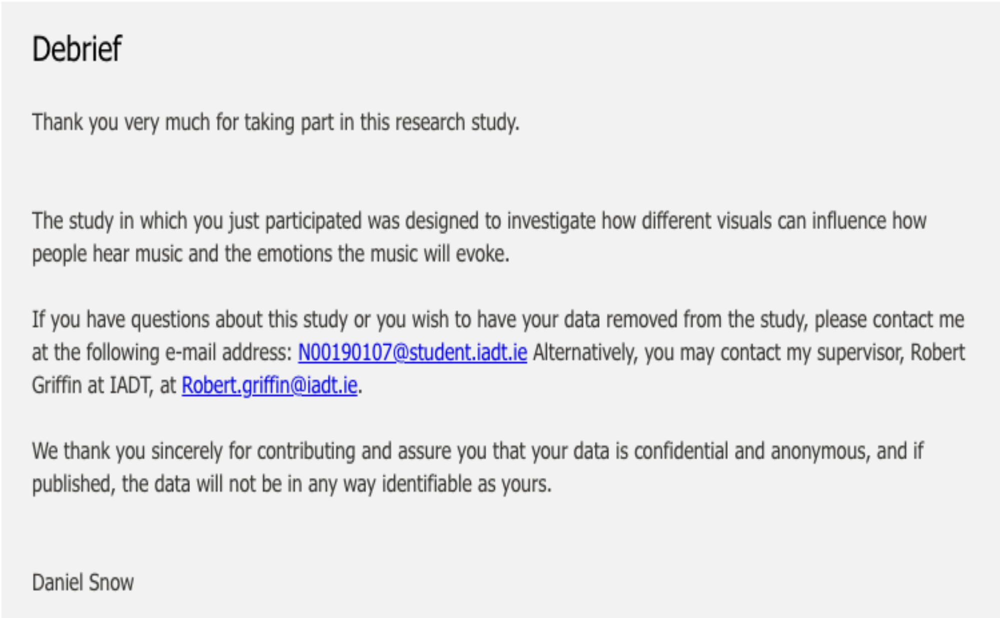
What if there is a problem?

If you are concerned about any aspect of this study, you may wish to speak to the researcher(s) who will do their best to answer your questions. You should contact Daniel Snow N00190107@student.iadt.ie or their supervisor Robert Griffin Robert.griffin@iadt.ie.

Thank you

Daniel Snow

Figure C-3
Debrief sheet



Debrief

Thank you very much for taking part in this research study.

The study in which you just participated was designed to investigate how different visuals can influence how people hear music and the emotions the music will evoke.

If you have questions about this study or you wish to have your data removed from the study, please contact me at the following e-mail address: N00190107@student.iadt.ie Alternatively, you may contact my supervisor, Robert Griffin at IADT, at Robert.griffin@iadt.ie.

We thank you sincerely for contributing and assure you that your data is confidential and anonymous, and if published, the data will not be in any way identifiable as yours.

Daniel Snow

Appendix D

SPSS Output

Table D-1
Sample

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Column1	135	1.00	3.00	1.9926	.85078
Valid N (listwise)	135				

Table D-2
Levels of independent variable

Column1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	VideoRepetition	49	36.3	36.3	36.3
	VideoIntended	38	28.1	28.1	64.4
	NoVideo	48	35.6	35.6	100.0
	Total	135	100.0	100.0	

First Factor Categories

Table D-3

Tension significance

total tension across Column1

Independent-Samples Kruskal-Wallis Test Summary

Total N	132
Test Statistic	1.657[a,b]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.437

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

Figure D-1

Tension distribution

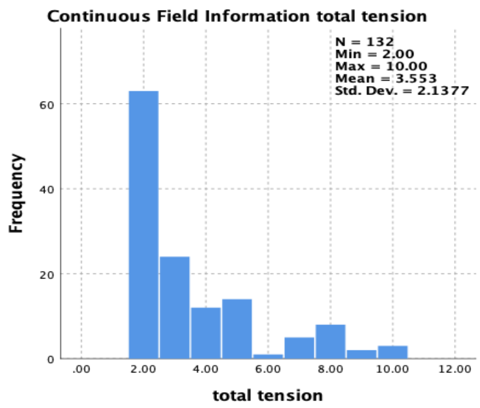


Table D-4

Power significance

total power across Column1

Independent-Samples Kruskal-Wallis Test Summary

Total N	133
Test Statistic	2.458[a,b]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.293

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

Figure D-2

Power distribution

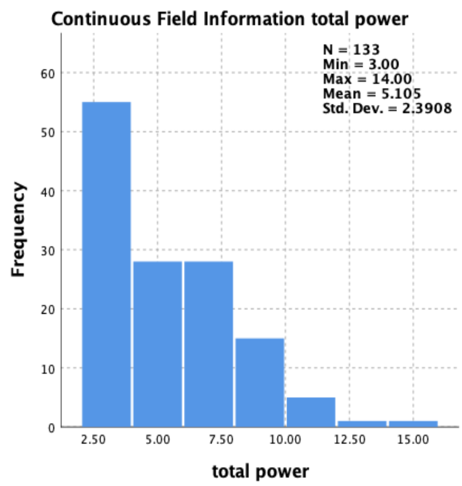


Table D-5

Transcendence significance

total transcendence across Column1

Independent-Samples Kruskal-Wallis Test Summary

Total N	128
Test Statistic	4.730[a,b]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.094

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

Figure D-3

Transcendence distribution

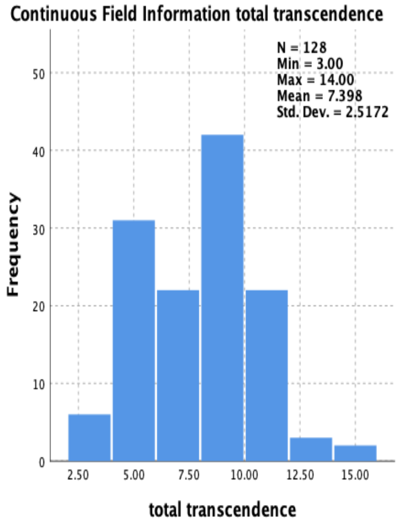


Table D-6

Wonder significance

total wonder across Column1

Independent-Samples Kruskal-Wallis Test Summary

Total N	131
Test Statistic	3.101[a,b]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.212

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

Figure D-4

Wonder distribution

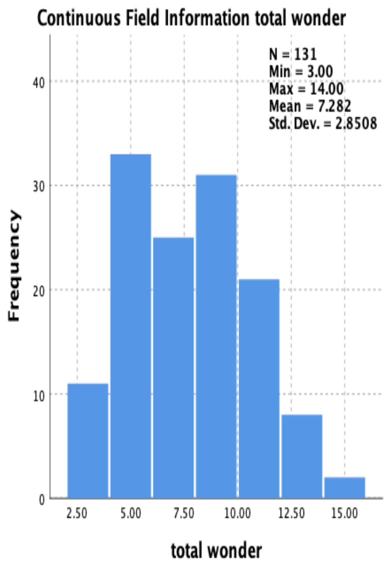


Table D-7

Joyful Activation significance

total joy act across Column1

Independent-Samples Kruskal-Wallis Test Summary

Total N	130
Test Statistic	.711[a,b]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.701

Figure D-5

Joyful Activation distribution

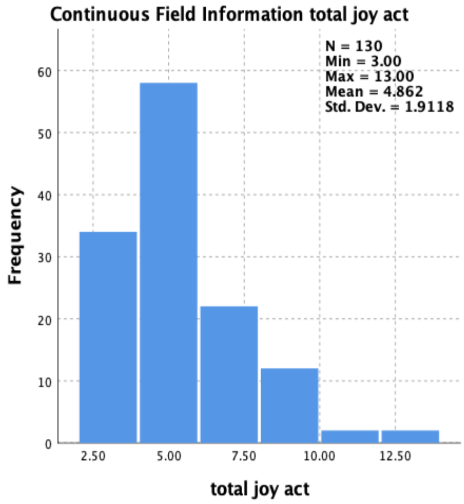


Table D-8

Significant result Peacefulness

total peacefullness across Column1

Independent-Samples Kruskal-Wallis Test Summary

Total N	132
Test Statistic	10.476[a]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.005

a. The test statistic is adjusted for ties.

Table D-9

Comparison of video conditions

Pairwise Comparisons of Column1

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
VideoRepetition-NoVideo	-19.052	7.770	-2.452	.014	.043
VideoRepetition-VideoIntended	-25.243	8.393	-3.008	.003	.008
NoVideo-VideoIntended	6.191	8.393	.738	.461	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple ...

Figure D-6

Peacefulness distribution

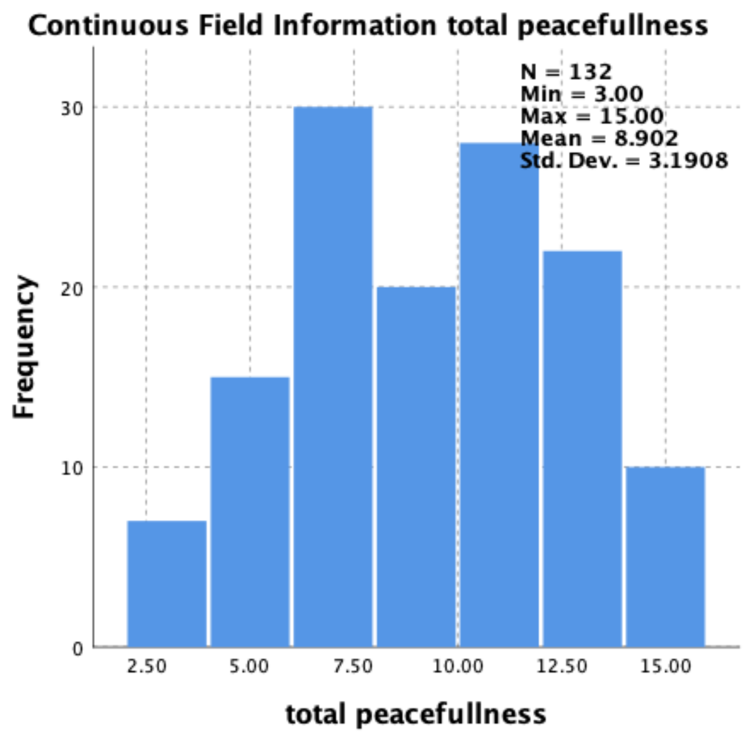


Figure D-7

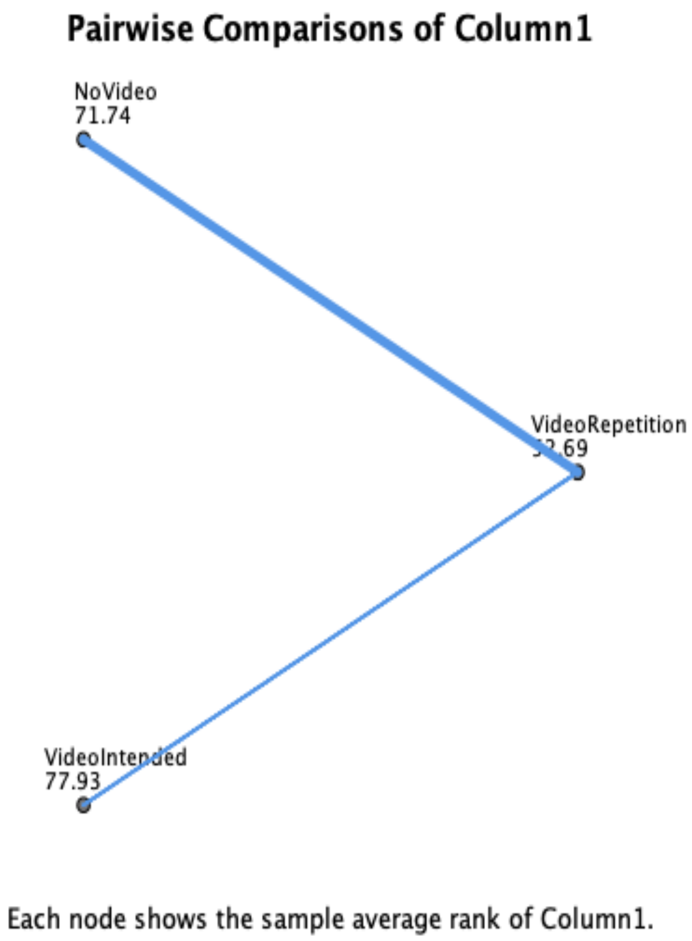


Table D-10

Median scores for Peacefulness three video conditions

Report

Median

Column1	total peacefullness
VideoRepetition	7.0000
VideoIntended	10.0000
NoVideo	9.5000
Total	9.0000

Table D-11
Nostalgia significance

total nostalgia across Column1

Independent-Samples Kruskal-Wallis Test Summary

Total N	131
Test Statistic	5.514[a,b]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.063

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

Figure D-8
Nostalgia distribution

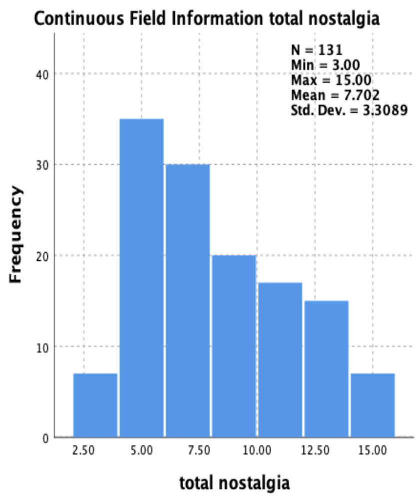


Table D-12

Tenderness significance

total tenderness across Column1

Independent-Samples Kruskal-Wallis Test Summary

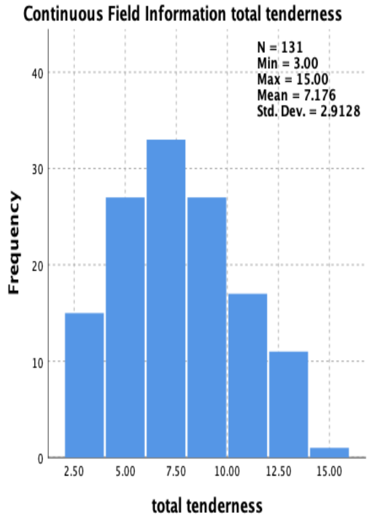
Total N	131
Test Statistic	4.091[a,b]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.129

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

Figure D-9

Tenderness distribution



Second Factor Categories

Table D-13

Sublimity significance

Sublimity across Column1

Independent-Samples Kruskal-Wallis Test Summary

Total N	124
Test Statistic	4.388[a,b]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.111

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

Figure D-10

Sublimity distribution

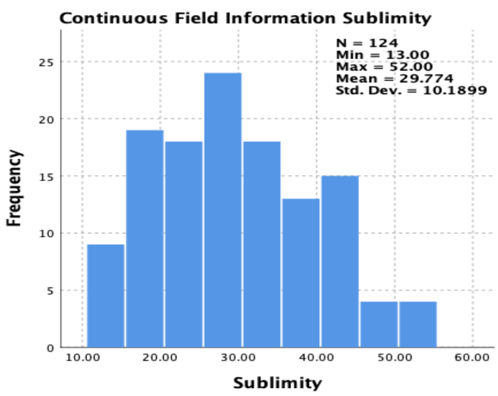


Table D-14

Figure D-11

Vitality significance

Vitality across Column1

Independent-Samples Kruskal-Wallis Test Summary

Total N	130
Test Statistic	1.539[a,b]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.463

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

Vitality distribution

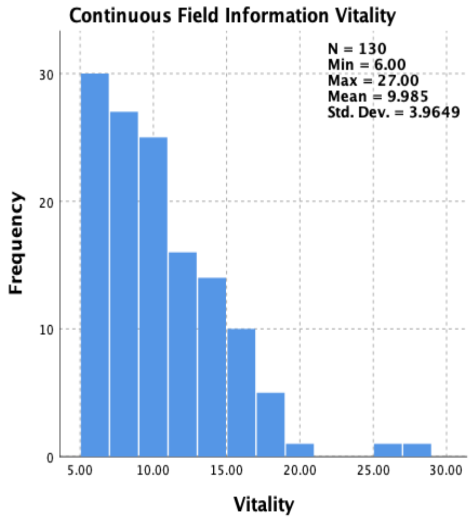


Table D-15

Unease significance

Unease across Column1

Independent-Samples Kruskal-Wallis Test Summary

Total N	132
Test Statistic	1.323[a,b]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.516

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

Figure D-12

Unease distribution

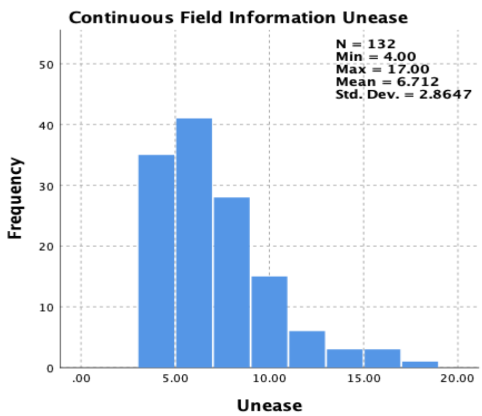


Table D-16
Affect significant result

Affect across Column1

Independent-Samples Kruskal-Wallis Test Summary

Total N	133
Test Statistic	6.262[a]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.044

a. The test statistic is adjusted for ties.

Table D-17
Comparison between groups

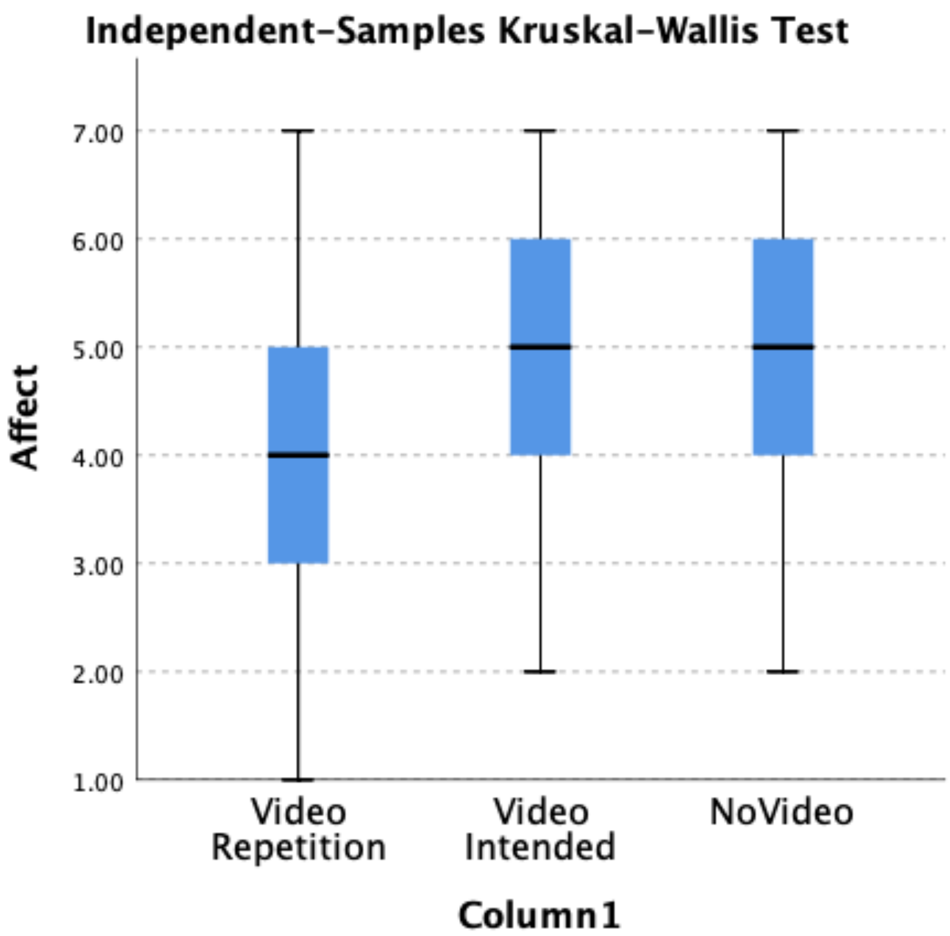
Pairwise Comparisons of Column1

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
VideoRepetition-NoVideo	-14.296	7.748	-1.845	.065	.195
VideoRepetition-VideoIntended	-19.323	8.198	-2.357	.018	.055
NoVideo-VideoIntended	5.027	8.237	.610	.542	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple ...

Figure D-13


Independent-Samples Kruskal-Wallis Test
Affect
7.00
6.00
5.00
4.00
3.00
2.00
1.00
Video Repetition
Video Intended
NoVideo
Column1

Perceptual scales

Table D-18

Activity significance

Activity across Column1

Independent-Samples Kruskal-Wallis Test Summary

Total N	135
Test Statistic	.331[a,b]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.847

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

Figure D-14

Activity distribution

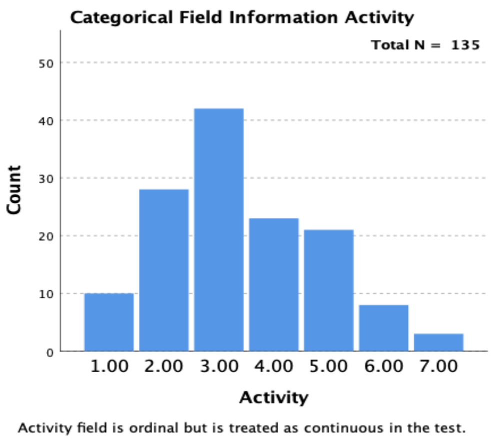


Table D-19

Tempo significance

Tempo across Column1

Independent-Samples Kruskal-Wallis Test Summary

Total N	135
Test Statistic	2.097[a,b]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.351

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

Figure D-15

Tempo distribution

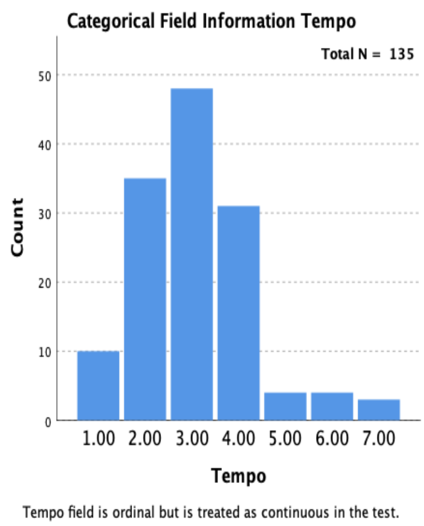


Table D-20

Rhythm significance

Rhythm across Column1

Independent-Samples Kruskal-Wallis Test Summary

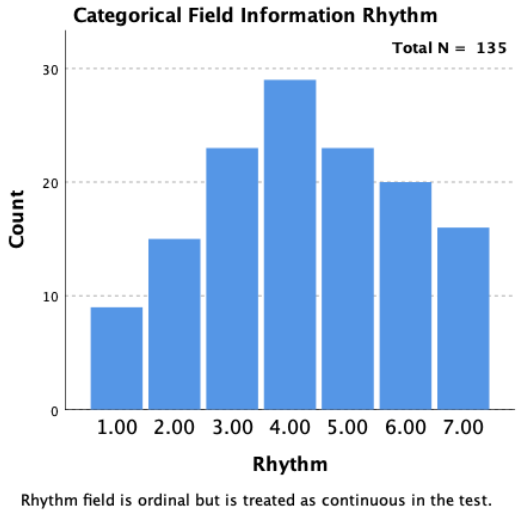
Total N	135
Test Statistic	3.425[a,b]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.180

a. The test statistic is adjusted for ties.

b. Multiple comparisons are not performed because the overall test does not show significant differences across samples.

Figure D-16

Rhythm distribution



Appendix E

Open ended questions transcripts

Original video

1	anonymous	Yes, both had dreamy , formless qualities
2	anonymous	Yes, the music was contemplative, which matched my curiosity about the visuals
3	anonymous	Yes as it matches the irregularity of the visuals
4	anonymous	Yes, vídeo movements fits well with the synth pads
5	anonymous	Very well matched as both shared a futuristic quality. They also both implied that the future was a calming and peaceful place / time. They each flowed together and felt congruent with one another.
6	anonymous	Yes, added to the sense of calm discovery as the video progressed
7	anonymous	Yes, ver serene overall, though also slightly ominous in places
8	anonymous	Yes, in that the mystery of the journey that the video was taking me on was well matched by the music that had a mysterious quality to it.
9	anonymous	Yes,the tempo and contemporary,synthesised sounds matched the futuristic visuals
10	anonymous	I agree, the music added to the surreal feeling of the video.
11	anonymous	They somewhat matched together. I could see it being a scene in a science fiction movie, but by itself it wasn't the best fit. The

		scene showed a fancy space ship medical leave place(?), but the music seemed mostly pretty calm and peaceful.
12	anonymous	Matched well: Calming colour palate , gentle pace , keeping attention by slowly refreshing the scene and music did same gently making minor alterations, both futuristic and other worldly
13	anonymous	Yes. The consistency of the music & its legato feel lined up with the spaciousness of the environment, & its many long unbroken lines. The music made me feel optimistic about the videos futurism more so than intimidated
14	anonymous	Yes reasonably well matched. The music was a repetitive chord sequence ?vi, ii, V, I, that had very little colour. The visuals were monochrome with short sequences but little stimulation.
15	anonymous	Somewhat agree that music and visuals were well matched. The spaces shown in the fly through were devoid of textures, human scale, human habitation. But there were still signs that this was a space used by humans, elevators, underground train, escalators, CT scan machine, rotating sleeping pods. Called to mind airports, train stations, hospitals. The music was universal, sounded like cultural influences in vocals, yet it wasn't specific to anywhere, like the spaces represented in the video. Lonely sounding like the empty spaces shown.
16	anonymous	Yes, the transition of the chords matched the movements of the scenery - I felt as I was in pace with both music/visuals. The vocals did not match as well and I found it a bit distracting.
17	anonymous	Steady visual pace of movement matched regular, cyclic chords, and the subliminal dissonant tones later made more emotional when brought into the vocal lent ominous /foreboding flavours

18	anonymous	Agree, calm music reflected the soft colour and landscape
19	anonymous	somewhat. the motion matched but feel of a surgical theatre or cryogenic lab not sure if that is the music? not sure?
20	anonymous	Yes
21	anonymous	agre
22	anonymous	I agree, mainly because of the tempo of the movement and the colour(s)
23	anonymous	Yes. The music was aerial and floating, like the vídeo.
24	anonymous	Yes.Both gave a sense of transcendence with a futuristic, and sci-fi feel.
25	anonymous	Yes, there is a sense of mystery caused by both
26	anonymous	Agrees if intent was to create serinity and a kind of wonder with the possibility of turning dark.
27	anonymous	Initially no, I thought the visuals far more fluid than the music but as they went on maybe a little more so.
28	anonymous	I agree with this. The music provided colour and evoked emotion from the detailed yet colourless video.
29	anonymous	Yes, The white visuals relate to the soothing music and gave an overall calming experience.
30	anonymous	Before entry in to the dark room the tone change was too early. You have suddenly very sinister music without the visual queue that would require it.
31	anonymous	Agree
32	anonymous	yes

33	anonymous	Yes I agree that they matched well. The visuals seemed futuristic and the music matched this well. The audio was dreamy, calming, and simplified. Both seemed to convert a feeling of Modernity
34	anonymous	Disagree - The video to me appeared to portray a futuristic looking hospital, but the music symbolised something slightly spiritual to me.
35	anonymous	Yes. Felt they were setting a scene without a conclusion
36	anonymous	They were both calm and somewhat repetitive but I don't have any strong feelings one way or the other
37	anonymous	Agree
38	anonymous	Agree

Repeated video

1	anonymous	Moderately well. The aesthetics worked with the music but the visual movements were a bit out of sync rhythmically.
2	anonymous	Yes. Something about the fast, smooth movement through the buildings and the slow, static nature of the music seemed to match.
3	anonymous	I think the they were. The shininess of the visuals, this kind of clean, futuristic aesthetic fut quite well with the sound that felt progressive but in a non-threatening way. The only way I can think of describing the relationship or the experience, is being *of the future*

4	anonymous	Yes, overall the music and visuals were well matched although towards the middle there was a slight change in the music (irregular additional string instrument sounds) which gave me a completely different feeling (almost the opposite - i.e tense and agitated) to the rest of the music. My earlier answers were based on the feelings experienced during the bulk of the video and music. The music gave a feeling of calm and and serenity which matched well with the clean, white, progressive and bright imagery.
5	anonymous	The less active moments enhanced the visual aspect. I found it conflicting to relate natural elements of the instrumentation to the digital aesthetic. The visual experience had a hurried pace to it, a sense of going nowhere fast, while the music felt restrained and open to exploration.
6	anonymous	Yes the gentle movement in the music matches the flow of the image s
7	anonymous	Yes, they matched well. Visuals and sound were clear and clean and there was some progression. More so in the music than the repeating visuals. There was regularity, evenness in both.
8	anonymous	The repetition of the visuals detracted from the relationship between the music and visuals. They were matched well before the looping, that is to say, the relationship changed once the visuals seemed to merely repeat.
9	anonymous	Yes surreal and patterned
10	anonymous	Yes - synths match the virtual environment
11	anonymous	The visual seem to me as being in a loop but every time getting faster. The music didn't reflect this and kept quite linear throughout
12	anonymous	No, the visuals were too bright, colorless and moved too fast. The first bars of the song felt more towards serenity than what the visuals did, (nice synth). The visuals invoked something disturbing, the music took too long to add those elements.
13	anonymous	Somewhat, although I felt the visuals were a little too quick for the music. The dream-like music suited the unusual visuals though.
14	anonymous	Yes. Repetitive music & video
15	anonymous	agree

16	anonymous	Yes .. both had a transcendent quality
17	anonymous	They felt as if they were matched great at the beginning. Then the visuals were repetitive whereas the music changed and had additional vocals/instruments?
18	anonymous	Very suited music, Its a mesmerising video and the music is as equally interesting
19	anonymous	Yes and no. The visuals were very neutral so it is easy for the music to influence the interpretation of the visuals, yet a very different music would have matched the visuals equally (since it it neutral)
20	anonymous	A part of me says yes, a part of me says no. On the fence with this one.
21	anonymous	Visually I would have liked to see more colour to match up with what my minds eye was seeing as a result of the sound scope
22	anonymous	The visuals got faster when the music got tenser. I think they were matching in that sense.
23	anonymous	A little mismatched, speed of the visuals seemed to be out of kilter with tempo of the music
24	anonymous	I think the visuals matched when it came to the alluring and almost mysterious nature of the music and the 3d space. However, there were some timing issues, the loop of the music and the video synced in terms of some predictability.
25	anonymous	Yes, this type of music you would use in a waiting room or elevator and in the video you were constantly waiting for something
26	anonymous	Agony
27	anonymous	no, the music made me feel calm and the visuals had the opposite effect
28	anonymous	Yes, the music had a futuristic vibe which suited the futuristic video.
29	anonymous	Yes
30	anonymous	Reasonably. Combining an ambient Eno-esque soundtrack with a fictional, repetative virtual reality is quite common. I

		was waiting for an electro beat coming in. Reminiscent of a computer game (e.g. Portal).
31	anonymous	No, I do not believe that the music and visuals were matching because because I found the music slow paced and the animations quite fast paced.
32	anonymous	Agree. The repeated loop worked as the music didn't develop a lot and didn't have a lot of tension and release. The visuals added to the feeling of staying in one place.
33	anonymous	I would agree. There's a sort of smoothness to both. In the video things feel plain, which is mirrored in the simple moving music. The only missmatch might be the video speed, if the visuals were slowed a little it might have merged somewhat better with the audio.
34	anonymous	Yes, a mixture of dreamy & cyber/digital sounds and visuals
35	anonymous	Throughout the video I increasingly felt the music and visuals were matched well. As time went on the rhythm of the music and the sequence of visual shots became more connected due to the repetition.
36	anonymous	Yes, but also that different music matched would have a whole different experience.
37	anonymous	Well matched in the sense that I can't imagine a different soundtrack, or a more appropriate one, to accompany the visuals. I'm inclined to think that the visuals lent the music a more sterile, less vivid quality.
38	anonymous	Yes
39	anonymous	The pristine, futuristic aesthetic matched the ethereal feel and instrumentation of the piece.
40	anonymous	disagree
41	anonymous	Yes, the movements in the visuals seemed to be timed to match the muscial piece.
42	anonymous	i agree
43	anonymous	No reaction. The construction and presentation of this as an experiment means that it is intended to illuminate something. That's all.

44	anonymous	I would say they matched well, the slow music seemed to fit the mostly colorless modern visual style.
45	anonymous	Yes calm
46	anonymous	Not well matched. The visuals were colder than parts of the music
47	anonymous	Both music and visuals repetitive and quite tedious
48	anonymous	No. Speed of graphics and music didn't match so well.
49	anonymous	Yes, sense of mystery and unknown in visual similar to the music

Table E-1
Total scores original video

Original video coding

	Coder 1	Coder 2	Author
Agree	30	29	31
Disagree	2	3	2
Neither	6	6	5

Agree =1. Disagree = 2. Neither agree or disagree = 3

Table E-2
Inter rater reliability original video

	Coder 1	Coder 2	Author	Agreement
ID 1	1	1	1	1
ID 2	1	1	1	1
ID 3	1	1	1	1
ID 4	1	1	1	1
ID 5	1	1	1	1
ID 6	1	1	1	1
ID 7	1	1	1	1

ID 8	1	1	1	1
ID 9	1	1	1	1
ID 10	1	1	1	1
ID 11	3	3	3	1
ID 12	1	1	1	1
ID 13	!	1	1	1
ID 14	3	3	3	1
ID 15	3	3	3	1
ID 16	1	1	1	1
ID 17	1	3	1	0
ID 18	1	1	1	1
ID 19	3	3	3	1
ID 20	1	1	1	1
ID 21	1	1	1	1
ID 22	1	1	1	1
ID 23	1	1	1	1
ID 24	1	1	1	1
ID 25	1	1	1	1
ID 26	3	1	1	0
ID 27	3	2	3	0
ID 28	1	1	1	1
ID 29	1	1	1	1
ID 30	2	2	2	1
ID 31	1	1	1	1
ID 32	1	1	1	1
ID 33	1	1	1	1
ID 34	2	2	2	1
ID 35	1	1	1	1
ID 36	1	3	1	0
ID 37	1	1	1	1
ID 38	1	1	1	1
Total =				34 (89.5%)

Table E-3

Total scores repeated video

Repeated video

	Coder 1	Coder 2	Author
Agree	27	29	28
Disagree	9	9	9
Neither	11	9	10

Table E-4
Inter rater reliability repeated video

	Coder 1	Coder 2	Author	Agreement
ID 1	3	3	3	1
ID 2	1	1	1	1
ID 3	3	1	1	0
ID 4	1	1	1	1
ID 5	3	2	2	0
ID 6	1	1	1	1
ID 7	1	1	1	1
ID 8	3	3	3	1
ID 9	1	1	1	1
ID 10	1	1	1	1
ID 11	2	2	2	1
ID 12	2	2	2	1
ID 13	3	3	3	1
ID 14	1	1	1	1
ID 15	1	1	1	1
ID 16	1	1	1	1
ID 17	3	3	3	1
ID 18	1	1	1	1
ID 19	3	3	3	1
ID 20	3	3	3	1
ID 21	2	3	3	0
ID 22	1	1	1	1
ID 23	2	2	2	1
ID 24	3	3	3	1
ID 25	1	1	1	1
ID 27	2	2	2	1
ID 28	1	1	1	1
ID 29	1	1	1	1
ID 30	3	3	3	1
ID 31	2	2	2	1
ID 32	1	1	1	1
ID 33	1	1	3	0
ID 34	1	1	1	1
ID 35	3	1	1	1
ID 36	1	1	1	1
ID 37	1	1	1	1
ID 38	1	1	1	1
ID 39	1	1	1	1
ID 40	2	2	2	1
ID 41	1	1	1	1
ID 42	1	1	1	1
ID 44	1	1	1	1
ID 45	1	1	1	1

ID 46	2	2	2	1
ID47	1	1	1	1
ID 48	2	2	2	1
ID 49	1	1	1	1
Total				43 (91%)