The Sound of Paintings: Using Citizen Curation to Explore the Cross-Modal Personalization of Museum Experiences

Paul Mulholland
Knowledge Media Institute, The Open University, Milton Keynes, UK
paul.mulholland@open.ac.uk

Adam Stoneman
Irish Museum of Modern Art, Royal Hospital Kilmainham, Dublin, Ireland
adam.stoneman@imma.ie

Naomi Barker
School of Arts & Humanities, The Open University, Milton Keynes, UK
naomi.barker@open.ac.uk

Mark Maguire
Irish Museum of Modern Art, Royal Hospital Kilmainham, Dublin, Ireland
mark.maguire@imma.ie

Jason Carvalho
Knowledge Media Institute, The Open University, Milton Keynes, UK
jason.carvalho@open.ac.uk

Enrico Daga
Knowledge Media Institute, The Open University, Milton Keynes, UK
enrico.daga@open.ac.uk

Paul Warren
Knowledge Media Institute, The Open University, Milton Keynes, UK
paul.warren@open.ac.uk

ABSTRACT
This paper describes the use of Citizen Curation to explore ways in which cross-modal experiences can be used and created by museum visitors. Citizen Curation can be defined as individuals and groups from outside the museum profession engaging in curatorial activities to communicate their own ideas and stories. Previous work has explored how Citizen Curation can be used to broaden the range of voices reflected in the museum, thereby widening its appeal and relevance to new audiences. Recent research suggests that cross-modal experiences, combining visual art with music, can enhance the cultural experience as the visitor simultaneously draws on both what they see and hear. Citizen Curation provides a potential method through which visitors can create and share cross-modal experiences for each other, combining visual art and music. In this paper, we introduce the Deep Viewpoints web application that has previously been used for the Citizen Curation of looking at visual art. We then describe how the application was extended to support two further contexts (i) a musicologist curating experiences that link music to visual art in a museum collection, and (ii) visitors to a museum exhibition experiencing and creating cross-modal experiences. Finally, we reflect on different ways in which technology could be used to support cross-modal museum experiences.

CCS CONCEPTS
• Human-centered computing → Human computer interaction (HCI); Empirical studies in HCI.

KEYWORDS
Museums, Personalization, Cross-Modal experiences, Citizen Curation

ACM Reference Format:

1 INTRODUCTION
The UK Warwick Commission report [40] highlights a huge disparity in rates of cultural engagement. Among museum visitors, 87% are from higher social groups. Similar trends are found in other forms of cultural participation. For example, the wealthiest, better educated and least ethnically diverse 8% of the population make up 28% of theatre audiences. Cultural participation among the BAME community and those with a disability is markedly lower. Cultural participation among the young has been on a downward trend across a range of the arts. However, some elements of culture are more widespread. 70% of adults consider music to be an important part of their lives. This increases to 84% among young people under the age of 25. The commission report suggests therefore that the relatively lower level of engagement among certain groups is not due to a general lack of interest in culture but rather a mismatch between their interests and what many public cultural institutions offer.

Our recent work has explored the use of Citizen Curation to widen cultural participation at the Irish Museum of Modern Art (IMMA) [49]. This has involved working with communities under-represented among museum visitors to develop and share their own perspectives on IMMA’s collection and exhibitions. These new perspectives on art can then be shared with other visitors, widening the range of voices included in the museum’s public offering and helping a broader range of potential visitors to connect art to their...
own lives and concerns. The activities at the museum in which the communities participated have made use of the process of Slow Looking [39, 44, 50] to help them engage with art and develop their own perspective. Slow Looking is a process that uses a series of prompts and questions to help the visitor to slow down and develop and share their own interpretation of the artwork.

Given the widespread passion for music, particularly among younger audiences who are underrepresented among museum visitors, this led to the question as to whether Citizen Curation, currently supported by a process of Slow Looking, could be expanded in which people from outside the museum could associate visual artworks with pieces of music, creating cultural experiences that involve not only looking but also listening. Combining music and visual art could potentially help museums to engage new audiences who would be able to develop and explore associations between visual art, with which they may be less familiar, and music, which is more likely to already play a significant role in their everyday cultural lives. As well as music being a potentially more familiar cultural form, research also suggests listening to music while looking at visual art can enhance the experience. Isaacoen et al [25] found that cross-modal experiences combining music and visual art, can help the observer to find new meaning and value.

The rest of this paper is structured as follows. Section 2 outlines related work on looking, listening and cross-modal experiences and how these can be used to personalise cultural engagement. Section 3 describes Deep Viewpoints, a web application to support Citizen Curation and how it was extended to support cross-modal experiences. Section 4 describes the use of Deep Viewpoints by a musicologist to create connections between a museum’s collection and pieces of music. Section 5 describes the use of Deep Viewpoints with a youth group, supporting them in both using and creating experiences that span music and visual art. Section 6 presents an analysis of the materials curated by the musicologist and the youth group. Section 7 discusses implications for this work in terms of how technology can be used to support cross-modal experiences in the museum.

2 RELATED WORK

2.1 Personalization and interpersonalization

Rydin and others [47] describe interpersonalization in which one museum visitor creates an experience for another. Interpersonalization is proposed as a complement or alternative to the use of algorithmic personalization to create museum experiences tailored to the needs or interests of the visitor [28]. Rydin and others [47] present two designs in which a museum visitor creates a museum experience for another visitor: Gift and Never let me go.

Gift [30, 48] is an app that visitors can use to create a gift for other visitors with whom they have a personal tie, such as a friend or relative. The gift takes the form of photographs of a small number of museum objects accompanied by a personal audio message. Never let me go [46] is an app designed as a two-player experience in which one museum visitor takes the role of controller and the other the role of avatar. The controller uses a menu of commands to communicate with the avatar. The available commands are organized into six different categories (such as body, questions and feelings). Selected commands (e.g. “Explore”, “Close your eyes”, “Who does it remind you of?”) are relayed to the avatar via audio, guiding their museum experience.

In this paper, we also explore how people can create museum experiences for each other. However, in our case we adopt the process of Citizen Curation, in which visitors curate new perspectives on the museum’s collection or exhibition. Unlike interpersonalization, the creator of the experience (the Citizen Curator) may not have any social tie to, or knowledge of, the later recipients of the experience.

2.2 Citizen Curation and Slow Looking

Mauer [32] and Hill et al [20] use the term Citizen Curation to describe citizens creating physical and digital exhibitions using archival materials from museums, libraries and other institutions. Citizens were trained in professional curatorial practice and applied their newly acquired skills to curate their own exhibitions. Such an approach requires a significant time investment over several weeks. In our own work [49] we explored how meaningful elements of curatorial practice could be undertaken by citizens with minimal training and requiring a much lower time commitment. This could significantly lower the participation barrier, perhaps particularly among groups that museums traditionally find hard to reach. As an alternative to a training program in curatorial practice, our work adopted Slow Looking as an easy-to-use method, requiring minimal training, that citizens could apply to develop and share their own responses to museum artworks.

Slow Looking uses prompts and questions to help the visitor to develop their own interpretation while recognizing the advantages of mindfulness and slowing down to make a personal connection to an artwork [39, 44, 50]. Slow Looking was already in use at IMMA in the form of Slow Art Videos [24], developed and presented by the museum’s Visitor Engagement Team (VET). In line with contemporary cultural practice, the videos not only provide the viewer with contextual information (for example, about the life of the artist, and the social and political context to which they were responding) but also ask questions (e.g. “What is our attention drawn to first?”) to prompt the viewer to think about what they see, similar to the use of Reflective Questions in museums [8, 43]. This process of guiding the visitor to form their own interpretation can be described as a form of mediation [38]. Producing videos, such as the Slow Art videos, to mediate the interpretation process, is complex and time-consuming. However, the Slow Looking method itself, realised in an alternative form, could potentially be used by visitors to curate and share their own perspectives.

Recent work has investigated the potential benefits of Slow Looking for the museum visitor. Chamberlain and Pepperell [10] discuss which characteristics of an artwork could be revealed by Slow Looking, such as contrast, spatial layout, and figure-ground segregation and why some artworks (e.g. those that are visually indeterminate) may particularly lend themselves to Slow Looking. Igdalova and Chamberlain [22] compare the effects of Slow Looking with different types of audio content (short introduction followed by silence, mindfulness guidance or historical information) and three types of image (photograph, representational or abstract) on mood change and aesthetic engagement. Mood improved across conditions. Aesthetic engagement was found to be highest when Slow Looking
was accompanied by audio providing historical information. Cotter et al. [12] compared the effects of mindful looking (a mindfulness exercise), curious looking (generating questions about the artwork) and unguided looking on changes in well-being, emotion and immersion. Well-being was found to improve across conditions. What someone may see when Slow Looking could potentially be influenced by prior familiarity with visual art. Not employing the Slow Looking method, Foreman-Wernet and Dervin [18] use Dervin’s Sense-Making Methodology (SMM) to compare how novices and experts look at art. They found that experts focus more on style and structure and novices more on emotional responses to the artwork. DeSantis and Housen [16] propose that art appreciation proceeds through stages of aesthetic development from the accountative (making concrete observations about the work) through to the interpretative (critical understanding of the work) and re-creative (linking the work to personal and universal concerns).

In our work, we investigate how Slow Looking can be used to support the process of Citizen Curation. In this paper we explore specifically how music could be incorporated into the process, in which citizens can curate associations between visual art and music and engage in both listening and looking when experiencing art.

2.3 Listening, emotion and mindfulness

Museums offer Slow Looking, or related methods such as Visual Thinking Strategies (VTS) [51], to help the visitor interpret visual art. Cultural institutions do not offer equivalent methods for music or other auditory content, probably due to the predominance of visual art, and perhaps also due to music listening being a more familiar experience. However, research has considered how people make sense of music.

Given the value people place on music as a way of evoking emotion [27] a significant body of research has investigated how emotion is expressed and evoked through music. Juslin and Laukka [26] investigate how emotion is conveyed through music by the composer and how those emotions are perceived by the listener. From a review of the literature, they observe that emotions are expressed using a combination of musical features in a probabilistic way. For example, anger may be expressed through a combination of features including fast tempo, minor mode, high sound level, small pitch variability and rhythmic changes. They note that a musical feature may be used in a similar manner in the expression of different emotions. For example, fast tempo can be used to communicate both anger and happiness. Emotions evoked in the listener may be rather different to those expressed by the composer, due in part to the social context in which music is consumed and the personal associations a listener makes between music and their lives. Hunter et al. [21] in their study of music and emotion distinguish between the listener’s perception of emotion in music (essentially identification of the composer’s intent) and the emotion they personally feel while listening. Although the two were correlated, perception ratings were more emotionally intense than feeling ratings.

Reybrouk [45] argues that musical sense-making is an active process in which the listener creates their own associations between music and meaning rather than necessarily decoding the meaning intended by the composer. One factor that can be expected to affect musical sense-making is the listener’s familiarity with the type of music being played. Davies [14] identifies the qualified listener as someone who is sufficiently familiar with a style of music to appreciate its expressiveness. Davies [14] notes that most qualified listeners do not have formal music education but have acquired knowledge through exposure.

Bodie and Jones [7] distinguish active listening and passive hearing, whether the source be music or another form of auditory information such as speech. Drawing on the dual process theory of reasoning [17] they note that even active listening combines both fast, intuitive, unconscious processing as well as slow, deliberative, conscious processing. A listener may have an immediate, intuitive response to a piece of music which they combine and rationalise with slower, conscious processing of the music. Bodie and Jones [7] note the relationship between being aware of these unconscious reasoning mechanisms and mindfulness [9] in which people try to access their unconscious reasoning and reflect on how it affects their thoughts and emotions. Similarly, mindfulness is a feature of Slow Looking in which the observer is prompted to consider not only what they see and how they feel, but also consider why, bringing their intuitive responses to conscious attention.

In our work, we explore how Citizen Curators could potentially combine Slow Looking with listening to enhance the cultural experience, responding to both visual and auditory information and reflecting why they respond the way that they do.

2.4 Cross-modal experience of audio and visual art

Previous work has investigated relationships between the perception of music and visual art. Limbert et al. [29] reported on a study in which participants observed a painting with either matching, non-matching or no music. Impressionist or abstract music was selected to either match or mis-match impressionist or abstract paintings. The aesthetic experience of viewing the paintings was found to be intensified when the paintings were accompanied with matching music. Palmer et al. [42] explored associations between music and color. Participants chose colors to fit different pieces of classical music. For both US and Mexican participants, faster major music was associated with saturated, lighter and yellower colours. Slower minor music was associated with desaturated, darker and bluer colours. A strong correlation was found between the emotions associated with matching colours and music, suggesting that music and colour association is mediated by emotion.

Albertazzi et al. [1] report on an experiment in which participants rated Kandinsky paintings and pieces of Schönberg music using Osgood semantic differential scales, rating the paintings and music as to extent to which they were, for example, fiery, warm or calm. They were also asked to match the paintings and pieces of music. Relationships were found between the ratings and the matches made between the paintings and pieces of music. Isaacson et al. [25] investigated how the perception of a painting can change when it is paired with music. Participants selected an artwork and then interpreted it, prompted by Slow Looking or visual thinking style questions (e.g. “What do you see?”). They then chose one of three musical pieces, pre-selected to match the painting, and repeated the interpretation exercise. The presence of music was found to enrich
the interpretation with, for example, a heightened perception of narrative and movement in the painting.

Benford et al [5] describe an emotional museum visitor experience called Sensitive Pictures. After choosing emotions associated with a painting on a web app, the visitor was instructed to locate it in the museum. Once they had found the painting, they played an audio dramatic fictional story offering an emotional interpretation of the painting. Visitors enjoyed the experience and attributed their enjoyment to the enhanced emotional response provoked by the painting and accompanying audio story.

In our work we aim to adopt an approach of Citizen Curation in which people from outside the museum create, share and experience associations between visual art and music. Prior work suggests cross-modal associations can enhance the museum experience, potentially by intensifying the aesthetic experience, enhancing visitor interpretation of the visual artwork in ways that might otherwise be overlooked, or increasing the emotional intensity of the experience.

3 DEEP VIEWPOINTS

Deep Viewpoints was co-designed with IMMA to support the process of Citizen Curation. Deep Viewpoints is used to experience and author Slow Looking-style experiences in the form of scripts which guide and sequence visitor interpretation across one or more artworks. Each script is made up of a set of stages. Each of the stages that comprise a script can be one of four types (figure 1). A statement stage provides information or a perspective and does not request any input from the follower of the script. A question stage asks a single question, for example about one or more artworks, to which the script follower can provide a response. A help text can also optionally be associated with the question which the script follower can choose to reveal for further assistance. A multi-question stage asks a set of questions about the same artworks. The author of the script can decide whether to present the questions in a predefined sequence (navigated by previous and next question buttons as in figure 1) or to be shuffled in a random order. The multi-question stage was proposed to parallel the onsite visitor experience in which a tour guide may ask a series of questions at a specific point in the tour. A multi-question stage has an optional help text that can provide additional guidance on how to answer the questions (e.g., look slowly at the artwork from left to right) or who the questions are for (e.g., these questions are for anyone who has emigrated to another country). A story stage provides a story stem or story opener and instructions on how to continue the story. This stage was inspired by the story completion method [33], a projective technique in counselling psychology for eliciting perceptions of the topic introduced in the story stem.

Responses created by following a script can be moderated by either the script author or administrator of the app. Moderated responses to any script can be accessed from a separate page within the app. Within the design, following a script and viewing the responses of others were deliberately separated. This design decision was taken to promote a divergence of responses, encouraging visitors to take part before viewing the responses of others.

Within Deep Viewpoints it is possible to respond to scripts either anonymously or logged into the application. If logged into the application, users maintain control over their contributions, and can delete or edit them at any time. An edited contribution that had previously been approved may return to the moderation queue for consideration by the script author. The application therefore provides each user control over their own contributions within the space of citizen curated content, whether authoring or responding to scripts [13].

Logged in users are also able to develop and share new scripts. A user can build their own personal selection of works by searching...
the IMMA collection (figure 2). The selected artworks can then be used in script authoring (figure 2). Each script can be associated with any number of the personally selected works. A script can be created as a sequence of stages. Zero or more artworks can be presented in each stage. Each stage type has custom fields for the types of text that it contains (e.g. questions, help text, story stem, etc.).

Deep Viewpoints was developed using the Angular web development framework. The citizen contributions and artwork metadata used within the app are stored in an underlying Linked Data infrastructure. Metadata of the IMMA collection was imported within the Linked Data infrastructure using SPARQL Anything [3] which can transform any structured data source (e.g. CSV, HTML) into RDF.

A musicologist (Naomi Barker, one of the authors of this paper) started to use Deep Viewpoints to create scripts that associated artworks, mainly from the IMMA collection, with pieces of music. For licensing reasons, associations were made between the artworks and recordings of the music found on YouTube rather than an alternative streaming service. Early in the process, additional requirements were elicited which motivated enhancements to the Deep Viewpoints app. First, support was added for embedding YouTube videos within the scripts to support listening concurrent with viewing artworks, reading associated text, and answering provided prompts within the page of the app. Second, functionality to support multiple choice as well as free text responses to questions was added to support the rating of music on a scale or selecting an emotion that matched the music. Third, a responsive web design (RWD) approach was taken to supporting both the following and authoring of scripts. This was done to potentially support the following and authoring of scripts on personal smartphones (potentially with headphones) while in the museum as well as on larger screen devices.

4 CROSS-MODAL EXPERIENCES WITH A MUSEUM COLLECTION

The enhanced version of Deep Viewpoints was used by the musicologist to develop 11 scripts associating artworks with pieces of music. The musicologist did not approach the task from the point of view of a specific methodology but drew on her own teaching practice and her research into the paintings of Poussin and his use of musical modes [4]. Some scripts were motivated by associations between the artwork title and the title and/or lyrics of the music as in the case of Eine Kleine Nacht Musick described below. In other cases, an association was made between what is depicted in the artwork and the music, for example, the tree depicted in the artwork The Music of Things (Sleep) mentioned below. In many cases, the musicologist drew on her teaching experience of using visual cues such as colour and shape to help students understand elements of musical structure and engage in close or analytical listening without knowledge of music notation. An analysis of the structure and textual content of the scripts (described later section 6) helped to reveal the fine-grained ways in which the scripts were being used to associate music and visual art. Below this is synthesized into the main types of association on the overall level of the script.

Art interpretation providing context for interpretation of related music. One of the scripts started by offering an interpretation of the artwork Il Trovatore by Giorgio de Chirico [15]. Explaining that the title translates as the troubadour, the script points to the long tradition of singing storytellers through to modern genres such as Rap. This set the scene for listening experiences with three varied pieces of music in the troubadour tradition.

Using emotion to associate visual art and music. Four of the scripts supported the reader in making associations between music and visual art based on emotion. For example, one script guided the reader in looking at the artwork Madonna and Child with Onlookers by David Godbold [19] and thinking about their
emotional reaction to the painting. They were then asked to select which of two pieces of music matched their response to the artwork.

**Imagining the look of a piece of music to make an association to visual art.** Three of the scripts asked the reader to imagine what a piece of music might look like and then related what they imagined to visual art. For example, one script contained the musical piece Walking the Dog by George Gershwin. The reader was invited to take a paper and pencil and see what marks they made while listening to the song. The script then moved onto looking at related visual art.

**Imagining the sound of an artwork to make an association to music.** Four of the scripts involved imagining the sound of an artwork and then making an association to music. For example, one script considered the artwork Eine Kleine Nacht Musik by Joseph Cornell and Robert Cornell [11]. The reader was asked to think about what sounds they associate with the night and also the cartoon-like mice in the artwork (figure 3, example 1). The script then moved on to consider the musical piece Eine Kleine Nachtmusik by Mozart.

**Imagining both the look of music and sound of an artwork to create two-way associations.** Three of the scripts involved two-way associations between music and visual art, that is, the three scripts that involved imagining the look of a piece of music also involved imagining the sound of an artwork. For example, one script started by looking at the artwork Ancient Music by Hughie O’Donoghue [41]. The reader was asked to imagine what sounds were evoked by the image. The script moved onto a listening experience with two pieces of music and in turn imagining how these pieces of music could be represented visually, essentially closing the loop from art to music and back to art.

**Synthesizing music into looking at visual art.** One script involved listening to different pieces of music while looking at the same artwork and reflecting on how the music changed the personal response to what was seen. This script focused on the artwork The Music of Things (Sleep) by Alice Maher [31], which appears to depict singing heads suspended from a tree. The reader was asked to consider how two pieces of music changed the perception of the artwork (Handel’s Ombra mai fu and Strange Fruit by Billie Holiday). This can be thought of as like the experiment of Isaacson et al [25] (see section 2) which investigated how music changed the perception of a painting.

**Synthesizing visual art into music listening.** One script involved synthesizing visual art into the listening experience. The script began by looking at View of Vetheuil by Clause Monet, which is a painting of a rural landscape. The reader was then asked to listen to an instrumental piece of music by Girolamo Frescobaldi,
chosen because it is based on a melodic pattern that represents the call of the cuckoo, and identify anything they heard in the music that reminded them of the countryside (figure 3, example 2). This can be thought of as the inverse of the experiment of Isaacson et al [25] (see section 2) in which visual art and music are combined but, in this case, primarily to change or enhance the listening rather than the looking experience.

As described in the next section, the scripts authored by the musicologist were later used in a cross-modal experience for a youth group visiting a museum exhibition.

5 CROSS-MODAL EXPERIENCES WITH A MUSEUM EXHIBITION

To investigate how cross-modal experiences could be integrated with a museum exhibition, an event was organized at IMMA for a youth group from the National Gallery of Ireland, called the Apollo Youth Panel. The event was structured in three parts.

The first part of the event explored looking and listening in the exhibition Kevin Mooney, Revenants [23], which, through a series of paintings, imagines the missing visual culture of Ireland lost to history due to poverty, famine and mass migration. Initially, the youth panel used the Slow Looking method to interpret the artwork Mutators by Kevin Mooney using typical questions such as “What do you notice first?” The panel then moved onto the artwork Ilcruthach by Kevin Mooney [35]. This time, interpretation was supported by three pieces of music played in the gallery while they looked at the artwork. Moody, dark and slow tempo music was selected to attempt to reflect the visual qualities of the painting. For the youth panel, Piano Piece 1952 by Morton Feldman drew attention to the isolation and mood of the figure in the painting. There was a shared feeling of sympathy for the figure being all alone in the world. The song Punisher by dubstep artist Pinch was associated with the inferred heartbeat of the figure, imagining the body of the figure as pulsating. It was suggested that the first piece of music drew attention to the head and what the figure was thinking and feeling, whereas this song brought attention to the torso and how the heartbeat might make it move. Finally, the song Te llevaré by Lisandro Meza was associated with the legs of the character and how it might sway or walk in a slightly stilted way in keeping with the music. It was commented how the pieces of music had moved them from the head to the heart and legs of the figure.

The panel then moved onto the artwork Blighters by Kevin Mooney [34] and its interpretation when accompanied by three further pieces of music. This time, music was selected on the basis of a conceptual connection to the subject of the painting, an imagined Irish diaspora. The three songs represented elements of diaspora or fusion between musical styles or cultures. With the song Djelem Djelem by Barcelona Gipsy Klezmer Orchestra, for the youth panel there was a focus on the mood of the song and a feeling that it reflected the sadness of the characters or their loneliness as the only survivors of some terrible event. The song Ka Bohaleng / On the Sharp Side by Abel Selaocoe evoked a more upbeat and optimistic interpretation. They felt that the song reflected the characters in the painting setting out on an adventure. They felt this could be thought of as the opening credits for a film about the characters as they are about to embark on their journey. It was commented how the music was helping to create a story around the artwork. Finally, the song The Wild Rover by Lankum was perceived as a sadder song more appropriate for the closing credits to a film reflecting the end of the journey of the characters in the painting. Now the mood of the characters did not seem very positive. It was felt that maybe the journey had ended, and the mission had failed. It was commented once again that the music was helping to create a story around the painting.

There were three striking aspects of the in-gallery cross-modal experience. First, the combination of music and visual art could influence interpretation in two very different ways. The three pieces of music accompanying Ilcruthach by Kevin Mooney [35] afforded a focus on parts of the character in the painting and the perception of movement. The three pieces of music accompanying Blighters by Kevin Mooney [34] helped to situate the characters in the artwork within a narrative. Second, individual pieces of music could lead to markedly different interpretations, evoking movement in different parts of the character (in the case of Ilcruthach) or different narrative extensions of the characters (in the case of Blighters). Third, these differing interpretations, whether evoking movement or narrative, could have an emotional element which could be either positive or negative.

The second part of the event took part in a studio away from the exhibition and involved following the scripts developed by the musicologist. The panel worked in small groups using either their own smartphones or tablet computers supplied by the museum. The youth panel contributed 22 rich and varied responses to scripts. For example, the prompt on how the artwork Ancient Music [41] would sound (mentioned in section 4) elicited responses including rain, slinky bugs, deep rumbling, organ and choir music and a musical piece that breaks to a softer tone. The artwork Eine Kleine Nacht Musick [11] (mentioned in section 4) was associated with many different sounds including jazz, techno, classical, the weather, birdsong, and a lullaby. The cartoon-like mice in the artwork raised several associations including children’s music, high-pitched, classical, piano, scurrying, jazz improv and squeaky music. The question of how to paint the musical piece Eine Kleine Nachtmusik by Mozart produced suggestions including zig zags, dancing on paper, swirling and rhythmic brushstrokes, vibrant and bright colours.

In the third part of the event the youth panel created their own cross-modal scripts in response to the Kevin Mooney, Revenants exhibition [23] visited earlier in the day. Working in groups and using a combination of the museum tablet computers and their own smartphones, they developed 6 scripts. They can be broadly organized into three groups.

Using music for art interpretation. Three of the scripts used music as an integral part of a response to one of the artworks. For example, one of these scripts made associations between contrasts found in the musical work An Ending (Ascent) by Brian Eno and those found in the painting Orbs [37] and other Kevin Mooney works (figure 3, example 3). Another of the three scripts used music to express something about a painting that could not easily be said. The script which included the artwork Ilcruthach by Kevin Mooney [35] (interpreted earlier in the gallery), contained the text “This strange character evokes a number of emotions and ideas. These cannot always be represented through words, but rather an
Imagining the sound of an artwork to make an association to music. Two of the scripts guided the follower of the script to think about sounds associated with the artwork as well as proposing their own musical association. One of the scripts was concerned with the artwork Island by Kevin Mooney [36]. The script asks the reader what environmental sounds they would hear in this location and what sounds the mountain depicted in the painting would make. The script also makes an association between the artwork and the song City made of code by Sungmin Park. These two scripts can be thought of as like the musicalogical scripts that encouraged the reader to imagine the sound of an artwork to make an association to a piece of music.

Using narrative to associate visual art and music. One script used narrative to make an association between visual art and music. Starting with the painting Blighters by Kevin Mooney [34] (interpreted earlier in the gallery), the script asked the reader a series of questions to build a narrative around the characters such as “Why are they here?”, “Who or what are they looking for?” and “What are they thinking about?” Later the script proposes the music Eyes Wide Open by Gotye and asks, “Does this song match your story?” (figure 3, example 4). This can be thought of as the inverse of the experience in the gallery, where pieces of music evoked a narrative. Here the script follower is encouraged to first develop a narrative and then consider a matching piece of music.

In the post-workshop discussion, the youth panel described the event as an immersive one, in which the music altered the viewing experience. They commented on how using the musicalogical scripts helped them to decide for themselves how they thought and felt about the visual artworks and music. When asked how the event could be improved or enhanced, they commented how the activity could be broadened further to include non-musical sounds and time for the drawing and painting of responses.

6 ANALYSIS OF THE CROSS-MODAL SCRIPTS

The classification of the cross-modal scripts described in sections 4 and 5 was guided by an analysis of their structure and thematic content as described in sections 6.1 and 6.2 below.

6.1 Script structure

The scripts developed by the musicologist and youth panel had a similar length, on average comprising approximately three stages: 3.27 stages for the musicologist scripts and 2.67 stages for the youth panel scripts (table 1). The scripts developed by both the musicologist and youth panel provided opportunities for the script follower to contribute a response. In each case, statement stages (that do not invite a response) made up just 14% of the scripts. The total number of prompts was calculated as summing each question within a multi-question stage with the number of story and question stages. Overall, the scripts offered on average at least one prompt per script stage (1.39 prompts per stage for the musicologist scripts and 1.00 prompt per stage for the youth panel scripts). This suggests that the script authors adopted a contemporary approach to curation in line with the Slow Looking method, encouraging the script follower to develop their own response as well as engage with the interpretations of the script author. This finding is substantiated by the thematic analysis of the script text in section 6.2.

6.2 Thematic analysis

The text contained in the scripts was classified in three ways. First, textual episodes were classified as to whether the script author was offering an interpretation to, or mediating an interpretation from, the script follower (table 2). In line with the analysis of script structure (section 6.1) numerous examples were found of mediation as well as interpretation. Second, a grounded theory approach [6] was adopted to further identify categories within the episodes of interpretation and mediation. Six thematic categories were found: observation; artist, personal, perspectival, societal and imaginative (table 3). This range of categories was found across both the musicologist and youth panel scripts. The themes reveal additional ways of responding that can potentially be used to create associations between art and music. For example, as well as the viewer or listener reflecting on their personal response, they can also consider what the artist intended or what characters depicted in the art or music might be thinking or feeling, as part of their response.

Finally, a grounded theory approach [6] was used to classify the episodes in terms of how they stayed within or transitioned between modalities (table 4). Examples were found that stayed within a single modality, that is, interpreting or mediating the looking at an artwork or the listening to a piece of music. The music episodes were found exclusively in the musicologist scripts. This may be due to the way the event was structured with the youth panel, with an initial focus on visual art. It may also reflect the musicologist’s greater familiarity with the description of music. Two types of cross-modal episode were found in the scripts: creating a cross-modal response (i.e. imagining the sound of a painting or the look of a piece of music) and creating a cross-modal association (i.e. associating an artwork and piece of music or asking for a response to a proposed association). As described in sections 4 and 5, cross-modal creation and association were often used in combination within the scripts, for example inviting the follower of a script to imagine what a painting might sound like and then asking whether a particular piece of music fits with what they are imagining.

7 DISCUSSION AND CONCLUSIONS

This work has explored how Citizen Curation can be adopted by people outside the museum to use and create cross-modal experiences. The in-gallery event with the youth panel and analysis of the scripts created by the youth panel and musicologist reveal three ways in which music and visual art can be combined in cross-modal experiences: scene-setting, transition and synthesis. First, one modality can be used to set the scene for engagement in the other modality. One of the scripts used an artwork (Il Trovatore by Giorgio de Chirico [15]) to introduce the concept of the troubadour before listening experiences with thematically related music. The inverse could also be anticipated, in which music is introduced to set the scene for looking at related visual artworks. Second, Citizen Curation can be used to make a transition or mapping from one modality to the other. As revealed in the script analysis, one technique that can be used is to first imagine what an artwork might
### Table 1: Analysis of the scripts developed by the musicologist and youth panel.

<table>
<thead>
<tr>
<th></th>
<th>Musicologist</th>
<th>Youth Panel</th>
</tr>
</thead>
<tbody>
<tr>
<td>N per script</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>N per stage</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

| Scripts        | 11           | 6           |
| Stages         | 36           | 16          |

<table>
<thead>
<tr>
<th>Type</th>
<th>N per script</th>
<th>N per stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statement</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Question</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td>Multiquestion</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>Story</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Prompts</td>
<td>50</td>
<td>21</td>
</tr>
</tbody>
</table>

### Table 2: Classifying episodes as interpretation and mediation.

<table>
<thead>
<tr>
<th>Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interpretation: script author offering an interpretation to the reader</td>
<td>In the seventeenth century, Giulio Mancini, a famous doctor wrote about how gazing at landscape paintings and imagining walking in the landscape depicted in them was as good for the health as walking in the countryside.</td>
</tr>
<tr>
<td>Mediation: script author assisting the reader to develop their own interpretation</td>
<td>Look at this painting. Using what you see as a starting point, what sensations do you imagine feeling if you walked in this landscape?</td>
</tr>
</tbody>
</table>

### Table 3: Types of interpretation or mediation carried out. All examples shown are mediation, except the societal example which is interpretation.

<table>
<thead>
<tr>
<th>Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation: an observation about the artwork or music</td>
<td>What are the first three words that come to mind when you look at this piece?</td>
</tr>
<tr>
<td>Artist: a view on the creator’s intent, work or background</td>
<td>The artist called this picture, The Sound of Silence. Do you think he found silence peaceful, scary or neutral nothingness?</td>
</tr>
<tr>
<td>Personal: a personal view on what the artwork or music makes you think or feel</td>
<td>Do you trust this bizarre character? Are you threatened or comforted by its presence?</td>
</tr>
<tr>
<td>Perspectival: someone else’s view such as that of a character in the artwork or music</td>
<td>Who or what are they looking for? What are they thinking about?</td>
</tr>
<tr>
<td>Societal: a societal and/or historical perspective related to the artwork or music</td>
<td>There are still singing storytellers today. There are a whole lot of modern genres that offer musical story telling about society – Blues, Rap etc.</td>
</tr>
<tr>
<td>Imaginative: an imaginative response to the artwork or music such as an answer to a creative question</td>
<td>Would you be scared by a giant spider? What would you do if it appeared next to you or if it bit you?</td>
</tr>
</tbody>
</table>

### Table 4: Modality of interpretation or mediation carried out.

<table>
<thead>
<tr>
<th>Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual art: Interpreting or mediating a piece of visual art</td>
<td>What do you see in this picture (lines, colours, spaces)? Do the patterns make you respond emotionally or affect your mood?</td>
</tr>
<tr>
<td>Music: Interpreting or mediating a piece of music</td>
<td>What story do you think they are singing? What aspect of the music strikes you first? How does the music make you feel?</td>
</tr>
<tr>
<td>Cross-modal creation: Creating a sound to match a piece of visual art or a visual to match a piece of music</td>
<td>What kind of music, if any, come to mind when looking at the painting? What sounds do you think the mountain would make? If you were to draw or paint the music, what would it look like? Does this song match your story?</td>
</tr>
<tr>
<td>Cross-modal association: Synthesizing or comparing a piece of visual art and a related piece of music</td>
<td>This strange character evokes a number of emotions and ideas. These cannot always be represented through words, but rather an eerie feeling within. [Text accompanied by music] Listen to these two songs focusing on the words. After listening to each one, look again at the artwork. Does the artwork have a different mood for you now?</td>
</tr>
</tbody>
</table>
sound like and then use that imagery to make and evaluate associations to pieces of music. Similarly, the scripts reveal transitions in the inverse direction, guiding people to express what they see while listening to the music and then using that to test and make associations to works of visual art. The scripts also demonstrated that both can be sequenced within a single experience, moving from one modality to the other before returning. Third, like the work of Isaacs et al. [25], music and visual art can be synthesized into a combined looking and listening experience. For example, the script focused on The Music of Things (Sleep) by Alice Maher [31] asked the reader to consider how two pieces of music changed perception of the artwork. Similarly, the synthesis of visual art and music can be used to change or enhance the listening rather than the looking experience. Work in one modality can also be used directly in an interpretative response to work in the other modality. For example, the youth panel scripts revealed examples of using music as an interpretive response to visual art which perhaps could not be described in words.

Our work also showed how different characteristics can be used as a “glue” when associating music and visual art: emotion, narrative and movement. Three of the scripts demonstrated how emotion can be used to connect music and visual art, reflecting on which pieces of music and visual art evoke a similar emotional response. This potentially connects to the work of Palmer et al. [42] (section 2.4) which suggests music and colour association is mediated by emotion. Narrative can be similarly used as a glue between music and visual art. The in-gallery event showed how music can help to evoke a narrative interpretation of art. One of the scripts created by the youth panel extended this idea by guiding the script follower to construct a story from an artwork which could then be compared against a piece of music. Movement can also be used in the association of visual art and music. During the in-gallery experience, the youth panel’s responses to Ilcruthach by Kevin Mooney [35] drew on the imagined movement of the figure in the painting, related to the rhythmic qualities of the music. Imagined narrative and movement used to connect music and art may also itself have an emotional element.

What does the cross-modal experience reveal that might otherwise have been missed? In line with, Isaacs et al. [25] our findings support the argument that music can help to heighten the perception of narrative and movement in visual art. Considering the inverse direction (visual art influencing listening) it could be argued that the perception of movement in visual art is drawing attention to the beat, tempo or rhythm of the music. Similarly, the narrative perception of visual art could be emphasizing the style or genre of the music. As well as potentially enhancing perception, feedback from the youth panel suggests that the cross-modal experience can be immersive. They also found the cross-modal scripts helped them to decide for themselves how they thought and felt.

The use of the Deep Viewpoints app has demonstrated how technology can be used to create not only Slow Looking experiences for each other, but also cross-modal experiences. Analysis of the script structure and content demonstrate that cross-modal scripts can be developed that not only offer an interpretation but also mediate the follower of the script in developing their own interpretation. Our work has adopted the approach of Citizen Curation in which experiences are created for others who may be unknown to the script author. However, cross-modality could also potentially be applied in the interpersonalization of experiences for friends or relatives.

Finally, the human curation or interpersonalisation of cross-modal experiences for others could potentially be assisted by algorithms, bringing human and computer personalization together in a complementary way. For example, algorithms could be used to help connect what a person thinks or feels in response to an artwork to candidate pieces of music that help communicate that feeling. The sentiment of the text produced while Slow Looking at an artwork could be matched to music which has lyrics of similar sentiment, or which has features that could be associated with the same emotion (e.g. the tempo, mode or rhythm).

Our future work will continue to explore how cross-modal experiences can be used to widen cultural engagement, helping people to make connections between art and their own concerns and interests. An element of this work may be the articulation of guidance for museum professionals on how to associate music and visual art to enrich the visitor experience. As suggested by the youth panel, this could also include associations to non-musical sounds, as in the use of sound-based micro-augmentations in the museum as proposed by Antoniou et al. [2] as well as associations between visual art and music.

ACKNOWLEDGMENTS

This work has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101004746 (Polifonia project) and grant agreement No 870811 (SPICE project).

REFERENCES

Examining the effects of slow looking on well-being in an online environment. Psychology of Aesthetics, Creativity, and the Arts.


[34] Adam Stoneman, Jason Carvalho, Enrico Daga, Mark Maguire, and Paul Mulholland. 2021. Uncomfortable Revelations: can citizen curation widen access to museums?. Museum Ireland, 28, pp.64-71.
