Nautilus – a case study in how a digital score transforms creativity.

Craig Vear, Carla Rees, Adam Stephenson

Abstract

In this article we discuss *Nautilus* (2022) a composition for solo bass flute, created using machine learning techniques and a Unity game engine. We discuss our approach and how this enhanced creativity and musicianship for those involved. We reflect on Unity's potential as a novel and flexible driver for the creation of a musical score, in which traditional elements of compositional design are presented to a performer as a co-creator for interpretation and communication inside the act of musicking. Through this, we offer insights into performer agency and the nature of decoding the media, sound, images, and AI presences within using his or her instrument, personal skills and musical aesthetic. Through the development process the notion of a music score was re-conceptualised and the activities of composition, collaboration and performance realisation were transformed. Although this is a single case study as part of a larger ERC funded research project (The Digital Score), the insights we discuss aim to be of benefit to those wishing to engage with digital scores.

Introduction

This article discusses a practice-based case study as part of an ERC funded project entitled *The Digital Score: Technological Transformations of the Music Score*[1] (DigiScore). The core aims of this frontier-research project are to: (1) determine scientific knowledge of how digital scores stimulate new creative opportunities and experiences within a range of music practices, (2) develop a theoretical framework for digital scores as an important transdisciplinary area of research, (3) build a scientific study of inclusive digital musicianship through the transformative potential of the digital score. This will be investigated through a series of practice-based case studies that place the experts at the centre of their practice so

as to offer them meaningful experiences with which to report back to the researchers of DigiScore who in-turn synthesise the results into a developing framework.

A key part of the framework that supports DigiScore is that meaning-making is to be understood from inside the create acts of music. To frame this approach, we adopt (evangelise even) Christopher Small's notion of *Musicking*[2]. In this he states that 'to music is to take part'[3], and that taking part can happen 'in any capacity, in a musical performance, whether by performing, by listening, by rehearsing or practicing, by providing material for performance (what we call composing)'[4]. Critically, Small stresses that 'the act of musicking establishes in the place where it is happening a set of relationships, and it is in those relationships that the meaning of the act lies'[5]. Simon Emmerson clarified Small's principle of 'meaning' to infer the 'what you mean to me'[6], (this subtle shift circumvents the significant issues of value and who is doing the evaluation of meaning). Therefore, meaning (or the *what-you-mean-to-me*) is to be found in the relationships formed between the new creative acts of musicking and the technologies and media of the digital score.

About Nautilus

Nautilus (2019-22) was the first case study of DigiScore. Although it started as a collaboration between the composer Craig Vear and low-flutes expert Carla Rees in 2019, it was stalled due to Covid and then picked up again at the start of DigiScore. For the purposes of this article, we will concentrate only on the latter development period.

The process brought together three practitioners to explore and contribute to the project from different perspectives – composer, digital score researcher and project PI Craig Vear, bass flute player Carla Rees and Unity programmer Adam Stephenson. Each brought their own unique experience and creative practice to the piece, driving both the aesthetic and practical considerations of the work. However, the core – and shared - goal was to create a digital score that supported and enhanced Carla's sense of musicking to a point that felt like it was operating *with her* in the making of the music.

On an artistic level *Nautilus* is inspired by an imaginary deep-sea journey of a nautilus mollusc, as it navigates deep water trenches across the oceans. The music describes this

journey with the bass flute and generative sound-design highlighting the topography of the oceans and vast openness of the depths.

<<< Figure 1. Screenshot of the visual part of the digital score for *Nautilus* >>>

On a technical level, *Nautilus* used a Unity game engine as the main platform for the visual elements of the digital score, which included sea-bed imagery and sinking notes written on staves tied to anchors (see fig. 1). The Unity engine also listens to Carla and makes judgements about what and when to generate a sound design.

The compositional process started with an improvisation by Carla on the idea of the nautilus' journey (the core aesthetic of the piece). This improvisation then became the source material for AI processes and the sound-design manipulation that is heard during the performance. The piece was developed using machine learning processes and used a neural net to make in-the-flow decisions about how the music is to be shaped. This was restricted to the development stage of the project and involved training a neural network[7] to predict and output a type of information based on input data. A neural network is basically a

complex statistical analyser used to output useful information from a given set of input data, for example to predict the house prices of a Boston suburb based on previous sale, size, demographic, socio-economic and location type data. For *Nautilus* the neural network was trained on transcribed jazz improvisations, so when a note extracted from her original improvisation was inputted, it calculated what the next note would be as if it was a jazz improvisor. In a music sense, it was trained to suggest new note possibilities through improvisation that expanded the scope of her original improvisation.

The machine learning process was the basis of a development environment (phase 1) but was eventually fixed and migrated into Unity behaviour (discussed below). At the start of each iteration of the development phase, random notes from Carla's original improv were passed through this neural network that in turn outputed a notated improvisation based on the input note choices. This notation formed part of the digital score for live interpretation and was designed to offer Carla a sense of familiarity (of her original improvisation) and suggestions of other materials (predicted by the neural network), so that she felt engaged in a shared/ co-creative musicking between AI and human musician.

Another element of the digital score is the generative sound design, which uses the audio recording of the original improvisation as its source material and responds to the live sound as a stimulus. These also manipulated the playback speed to give Carla a sense of familiarity and suggestion, and to further offer in-music relationships and a sense of musicking involvement. This was a way of drawing Carla into the complex programming of the AI, which can feel like a black-box devoid of musicking soul.

The final version of the digital score migrated all the learnt behaviour of the AI and neural networks into a single Unity environment. This creates an immersive world for the audience and musician to inhabit through the piece. This Unity engine version still generates random notation from a fixed library developed using the development processes, and also listens to Carla and generates a backing track using procedural algorithms designed to mimic the behaviour of the developmental AI. The Unity engine listens to the performer and its own backing track and moves the camera through the ocean using the amplitude of each source: left for live sound, right for generated sound. The aim was that through the developmental

phase Carla's regular engagement with, and incorporation of, the AI into her sense of musicking would become part of the way she responded to the computational elements as a co-operative other. By having traces of them in the Unity engine the hope was that she still felt their presence even if they were no longer AI generated. This was an unfortunate side effect of unifying all the programming into the Unity engine, which was a necessity of time and publishing limits.

It is crucial to state that the digital score consists of all these elements: Unity design, sinking notes, generative backing track, behaviour of the Unity "AI", the presences of thinking processes and audio files etc. Considering only the notation as the music score ignores the serious influence of all the other elements on meaning-making and what ideas are being communicated through the digital score, and ultimately how musicking is shaped by these elements.

This however presents a significant challenge to existing notions of musicianship: how to create a musicking experience that binds these elements together, rather than presenting something that is created from individual elements, like a Frankenstein score. The least successful version of this mixed-media approach would be to construct an experience for Carla that feels like a "bit of this", stuck together with a "bit of that". This unifying desire is immensely difficult as the creative musicians (coder, composer, performer), must work together to build something that has a purity and singularity of "message"[8], draws together and enhances the communicative value of a digital score, and has a unified aesthetic. This requires us to embrace a trans-disciplinary approach where we seek to find new, common principles and factors that contribute to a wholeness of all our musicking experiences, and this will normally go beyond/ distort/ transform/ enhance/ transcend our own training and ways-of-thinking.

For reference Carla's performance of *Nautilus* can be heard here
https://www.youtube.com/watch?v=XK-9eXCJxCg

Research Process

At the start of the research process, we identified key research questions that we wished to address through the practice as a way of seeking new insights, knowledge and understanding that contributed to the DigiScore investigation. These were:

- *can a games engine such as Unity be integrated into a digital score without detracting from the flow of musicking experience?*
- *how can a neural network be trained with an aesthetic design to generate a series of digital scores?*
- *how do we publish such a digital score so that others may engage with this composition?*
- *how can narratives structures from game design be used to enhance the experience of a musician with a digital score?*
- *performer agency in the interpretation of the score – how can this approach develop or enhance (or restrict) that further? and what are the challenges that arise from this?*

The research process was conducted in 3 phases:

1. *AI behavioural design and development* – here the behaviour of the AI and generative processes were designed and tweaked so that they felt part of the composition, not some extra process. This involved training and tuning neural networks within contained compositional aesthetics so that Carla was able to feel the presence of their behaviour and evaluate their affect upon her musicianship. These were relatively short examples that compositionally didn't progress but did allow the creative team an opportunity to grow the dynamic behaviours of the technology in line with the developing awareness of the compositional potentials of the piece.

2. *Compositional development* – once some of the parameters of the aesthetics and behaviours of the piece were understood, the composition was then allowed to develop. In a sense, the team had gained knowledge of these elements to such an extent that these relationships became the working materials for composition. Furthermore, the trans-disciplinarity of the creative team all had experience of these

materials and their potential meanings, so as to work together in the compositional process equally.

3. *Performance in a real-world environment* – the final test of *Nautilus* was to present it to a critical audience, and to conduct the qualitative data gathering from the creative team. This was designed using a two-way perspective of *encoding – decoding* where the experiences of those encoding communicative elements into a digital score (the composer, coder, designer) are compared with those experiences of those realising the digital score (performer, audience). Furthermore, the reflections of the composer, coder, designer are captured and included in this dataset, as is the legacy of experience on the performer through questionnaires many weeks after the event.

Reflecting on the meaning-making through musicking

In this section the team reflect on the meaning-making through musicking that happened through the research process. This is told through individual perspectives and is personal to each. They were generated by each completing an online reflection form as part of the dataset study and were conducted in isolation. As such, some perspective align, others don't. Overall, it presents a qualitative picture of the legacy of working on the project.

Performer's perspective – Carla Rees

The Unity immersive score gives the performer considerable agency in the interpretation of meaning through musicking. In a conventional score, pitches and rhythms are notated in detail, and space is left open for interpretation of nuance, taking into account relevant performance practices. In *Nautilus*, in addition to communicating these nuances, there is scope to develop an individual performance practice centred upon digital musicianship and musicking creativity. Visual objects propose relationships which are translated into sound, with decisions made according to personal aesthetics, visual awareness and the performance experience, since the material played changes the direction of travel through the virtual space through interaction with the audio file. The performer must notice and interpret a range of visual and aural cues in order to create a meaningful musical experience.

Choices can be made about which of the visual elements to play (since there are too many to play them all), as well as the durations of each sonic event. For example, a pitched note could last for as long as it can be seen on screen, but it may be musically more appropriate to move one's attention to another visual object long before it leaves the field of view. Everything seen (and heard) within the score provides potential for sonic material, and it is inevitable that different performers (and on different instruments) would decode these according to their own musicianship skills, technical toolkits, instrumental resources, and musical aesthetics.

The overall environment, combining visual and audio elements, however, provides information relating to the mood, atmosphere and general ambience which defines the overall character of the piece. The changing scenes can be interpreted in different ways, but the pace and energy of each scene defines the overarching structure of the composition and the underlying narrative. The piece can therefore be interpreted and decoded individually by each performer, with a certain level of creative freedom, while the piece itself maintains its overall identity. This is a very different way of interpreting and working with a score, and one that can shift the musicking experience for composers, performers, and audience.

Composer/ creative director perspective – Craig Vear

I felt that the choice of materials and behavioural presences were well judged for this composition. The iterative process helped immensely in identifying the value of each of these, and the open exploration process enabled radical thought and novel experimentation. It would have been great to explore the machine learning for longer, but we had arrived at a point were what we had was working, so it didn't feel necessary to investigate further. However, from a research perspective this core research question remains un-answered.

We created a fixed, but flexible 14-16 minutes piece with a linear sequence of form. Subsequently we adapted this to a single movement 6-minute version. This works well and guides (perhaps collaborates) with the performer in the construction of the music. We could have developed the code further to listen to her for cue's at transition points such as "wait for long held low C", but the migration to a Unity only system didn't easily support this

option, especially in the time constraints of the project. An open question remains about open-world potentials of this piece, removing the fixed form, or linear (progression) sequentially and entering a more emergent open-world format. One open question is about remote multi-player involvement in a digital score, like the experience of gamers in Fortnight. Bringing this together as a new case study with the same team is an exciting proposition.

It was interesting to feel how my role became more about creative director than a composer. In fact, the term "composer" was too reductive considering the roles and conversations that I needed to take on in order to drive the vision of this project forward and to accommodate the team's great ideas and input. Essentially, it felt like an auteur role but also like a parent who wished and welcomed ownership from the team members and relied on incorporating (or at least evaluating) all the great ideas that emerged. What was especially rewarding from this opportunity, was the enhancement that this approach offered the performer. I got the impression that so much more was being communicated through the relationships and presences that were "alive" inside the musicking experience. This offered many new ways of communicating to the performers and expanded the types of musical ideas that could be contained within a score paradigm.

Unity developer's perspective – Adam Stephenson

The interaction design in *Nautilus* was intended to give the performer a feeling of influence on the world's behaviours but not a feeling of exact control. The interaction system is inherently ambiguous, as is the backing track generation. It can decide to take input and react to the performer or ignore them and make its own decisions. These design choices were made to keep the performer from feeling as though they were in a game and could predictably control the outcome with certain clear behaviours. Instead, the performer may begin to recognise patterns and learn to adapt to the score, living alongside it and either playing with or against it. There is no hierarchy of control between the performer and the score, the two co-exist in the same space.

The presentation of the music notes in the scene as objects influenced by physics and instantiated with random velocities meant that sometimes the performer couldn't read and

interpret the note in time. They would also sometimes spawn too far away and couldn't be read. This could frustrate the performer, but it also feeds into the idea that this is a living world that doesn't make any efforts to accommodate them. The performer was also encouraged to make decisions and only play notes when and how it felt right for them in their on-going journey through the world. If I were to make any changes, I would increase the influence the performance had on the world. There were plans during development to add further interactions, but unexpected problems in other areas got in the way.

Developing with a specific user in mind that can interact with the work in a way that I cannot was a new challenge to me. In testing I could only emulate the performers input with taps on the mic and simple vocalisations, there was no way to see what effect a live flute performance would have outside of the few testing sessions. This led to me developing the systems to be easily, quickly, and deeply customisable to maximise the opportunity to get live feedback and implement changes during the few test sessions with the performer. This project taught me a lot about creating an effective project architecture which greatly speeds up incremental, experimental development.

The experience of working with experts in fields entirely different to mine has given me a greater perspective on collaboration. Each test session with the performer teased out so much more potential in the project due to their entirely different perspective on the score. My perspective was more technical and after seeing the visual presentation hundreds of times during development I'd grown somewhat insensitive to it. However, when the performer tested it, I was completely mesmerised and forgot about the hours spent tweaking code and creating the environment.

Audience perspective

As part of the research process, we conducted a survey of audience experience. Among general demographic data we asked two key questions pertinent to their reception of the music:

1. What musical ideas did the digital score communicate to you?
2. Do you feel you have learned something new about digital media and digital scores from this performance?

From an audience of 8 made up of master students on the MA/MSc in Creative Technologies at De Montfort University, we received 6 responses. The notable responses to these were:

"The 'scene' changes and differing in speed of movement always kept me engaged."

"It showed how a performer thinks throughout a performance but also how creative computing can influence their thought process. Seeing the environment change had an effect on my personal experience with the performance. I was looking both at the performer's reaction and my own trying to compare whether we share the same experience."

"It was interesting to see how a digital music score can be created for the musician to encode themselves and interpret in their own way, creating a unique piece of music each time."

"using instrument with Unity is amazing, it's inspired me to work music with digital technologies"

"the visual aspect of the score definitely influenced the performance, and allowed the performer to harness their pure reactions to the environment - resulting in a highly expressive performance."

"Music can communicate with environment created by Unity, new experience and performance I never seen before."

"It seemed to be more of an improvisational live scoring of a short film"
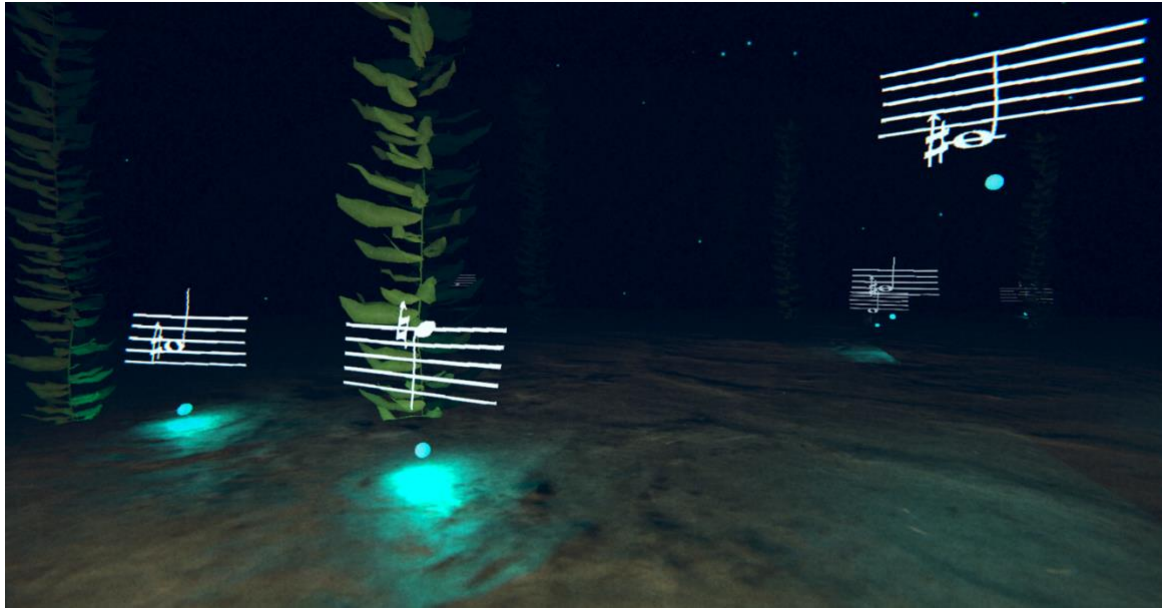
<<< FIGURE 2. Detail of flashing lights anchoring the notation to the seabed.>>>

Our findings to the key research questions

The section will discuss the implications of musicking with *Nautilus* using the key research of this case study questions as a framework (introduced above).

- *Can a games engine such as Unity be integrated into a digital score without detracting from the flow of musicking experience?*

Our findings, at least in this case study, is that Unity is an excellent environment for encoding and decoding musical ideas. The environment (visuals and sound design) provided a sense of mood/atmosphere which can be interpreted by the individual performer, as well as by the audience. The environment is immersive; and if carefully considered doesn't get in the way of the musicking experience. The question is how much to incorporate conventional notation and how much to leave free for improvisation? If one considers the environment to be similar to performance directions (tempo, mood, expression marks etc.) there is potential for pitch/rhythm information to be incorporated in any number of different levels of detail. For example, we used flashing lights to indicate rhythmic information, speed and density of notes dropping to invite interpretations about melodic line construction and gestures (see figure 2). There is also potential for this format to increase accessibility, as it could be created in a form that does not require an ability to read complex musical notation. This might make the music

approachable for performers from different musical cultures/backgrounds for whom historically European notation systems are not part of their working methods.

- *how can a neural network be trained with an aesthetic design to generate a series of digital scores?*

In *Nautilus* neural networks were used as a rapid-prototyping tool. Their role was to generate materials and provoke responses that aided in the development of the team's understanding of the potentials of this piece. In a sense they helped up find the boundaries of the composition and aesthetic and to test out behaviours from inside musicking. The use of transcribed jazz improvisations as training data was problematic and was felt to be a limiting factor. This is because jazz transcriptions are different from the aesthetic of Carla's own original improvisation and eclipsed it from a harmonic perspective. She felt that the jazz dataset was too limited harmonically and felt stuck in a language that was not her own or designed specifically for the bass flute. For her, it highlighted the need for the musical material and the visual environment to match in terms of potential for exploration. It would be interesting to explore how it might work using multiple versions of her own improvisation to create the dataset, in order to produce material that was more idiomatic for the Kingma system bass flute.

- *how do we publish such a digital score so that others may engage with this composition?*

Through discussions we arrived at a decision that publishing this piece needs to be in a format which is relatively inexpensive to produce and distribute, as well as simple for performers to set up without needing access to specialist software. Here Unity provided an interesting solution as it allows app exporting in both windows and Mac formats as well as through WebGL which means it could be hosted online and accessible through a webpage. This was one of the main factors for unifying all the processing into the Unity engine, rather than having separate systems for audio production, music notation generation, and visual design.[9]

>>> FIGURE 3. Detail of the *Nautilus* digital score showing section 3 "upside down world">>>

- *how can narrative structures from game design be used to enhance the experience of a musician with a digital score?*

Our approach allows for quick communication of ideas through the different worlds/structures/ universes which can break away from conventional 'rules' (eg gravity, going upside down; see fig. 3). This promotes creativity in terms of the performer having to choose how to interpret the various elements seen on screen, perhaps also breaking the musical 'rules' of a particular aesthetic approach. We found that the visuals can create emotions which can then be reflected in the music. The lack of/reduction in notated material invites the performer to develop memory skills in order to provide some structural coherence (e.g. remembering that particular symbols are performed in a particular way), but this is not significantly different from the requirements of improvisation or playing graphic or text scores.

- *performer agency in the interpretation of the score – how can this approach develop or enhance (or restrict) that further? and what are the challenges that arise from this?*

One finding with this case study, was that the performer can assign meaning to different objects most of which are not traditional or typical music notations. Furthermore, while different players might have the same sense of mood and even be using a similar "box of tricks" appropriate to the instrument and performance techniques, there is an opportunity to take an individual approach because of the way these materials invite a different mode or "message" in their musical meaning. This would mean that different performances could sound quite different but still maintain a form of broad aesthetic identity to the piece. In fact, following Carla's performance, the singer Franziska Baumann expressed an interest in realising *Nautilus*. Her performance can be heard here https://www.youtube.com/watch?v=SV6TqzJkiX4 and while it's a different instrument, and her interpretation varies from Carla's, the core aesthetic of the composition is retained.

Reflection

Working with digital scores in the way outlined in this article allows the performer to co-create the material and to bring their musical personality fully into the performance process. Traditional musicianship skills, such as listening, responding, and communicating ideas through one's instrument come to the fore, enabling the performer to become truly 'in the moment' and to experience a state of flow through the immersive nature of the materials. It is important that the music produced is in keeping with the ambience and mood created by the visual materials and soundtrack; without this the inherent logic of the narrative would be lost and the communication with an audience would risk breaking down.

Through this case study we have been able to begin to explore Unity's potential as a novel and flexible driver for the creation of a musical score. Traditional elements of compositional design (structure, narrative, mood, atmosphere) are presented to a performer for interpretation and communication, who engages with the materials using musical, technical and interpretative skills. The performer has agency in terms of developing the exact nature of the decoding of the visual cues, depending on his or her instrument, personal skills and musical aesthetic. An accompanying audio soundtrack also contributes to the musical direction of the work, enabling the performer to interact with sonic - as well as visual - cues

to help them to develop their own approach to the performance. This project has provided us with a springboard for potential future development, for example through interactivity with other performers.

Biographies

Craig Vear is Research Professor at De Montfort University where he is a director of the Creative AI and Robotics Lab in the Institute of Creative Technologies. He has been engaged in practice-based research with emerging technologies for nearly three decades, and was editor for *The Routledge International Handbook of Practice-Based Research*, published in 2022. His recent monograph *The Digital Score: Creativity, Musicianship and Innovation*, was published by Routledge in 2019, and he is Series Editor of Springer's Cultural Computing Series. In 2021 he was awarded a €2Million ERC Consolidator Grant to continue to develop his Digital Score research.

Carla Rees is a performer (low flutes, Kingma system and baroque flutes), arranger and composer. Her career focusses on collaboration and developing the dialogue between composer, performer and flute maker in order to extend and enhance the repertoire. She has premiered several hundred works in the UK and internationally. She is Professor of Low Flutes and Contemporary Flute at the Royal Academy of Music, runs an innovative distance-learning music degree programme at the Open College of Arts, and teaches the flute at Royal Holloway, University of London. Her works are published by Tetractys Publishing. www.carlarees.co.uk

Adam Stephenson is a freelance Creative Technologist and Games Programmer with a focus on interactive, immersive experiences through the lens of traditional games interaction design. His specialty is in 3D development using the Unity game engine to create Augmented and Virtual reality experiences, as well as traditional apps and games.

---

[1] Details available online at https://cordis.europa.eu/project/id/101002086 (accessed 25, May 2022)

[2] Small, Christopher *Musicking*. (Middleton, CT: Wesleyan University Press, 1989)

[3] Ibid, p. 9.

[4] Ibid.

[5] Ibid. p. 13

[6] Emmerson, Simon *Living Electronic Music*, (London: Routledge 2007), p. 29

[7] The training code used this repository with some minor changes
https://github.com/haryoa/note_music_generator/blob/master/Music%20Generator.ipynb

[8] Discussed in context online at https://digiscore.dmu.ac.uk/2022/01/17/the-digital-score-through-the-medium-and-its-message/ (accessed 25, May 2022)

[9] Mac and Windows versions of the digital score can be downloaded at
https://digiscore.dmu.ac.uk/2022/01/27/nautilus/