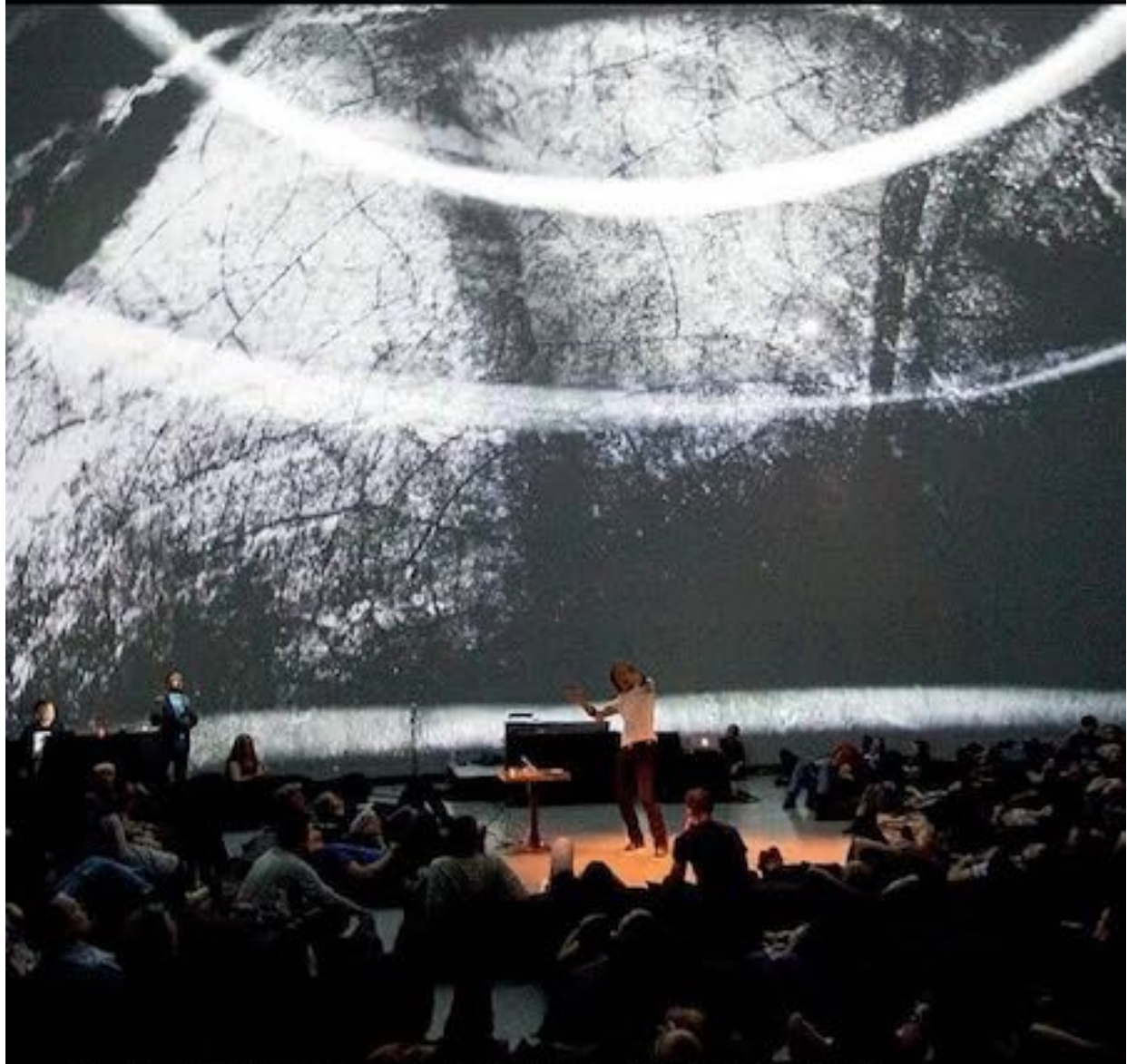


Sá, Adriana, & Antunes, Rui Filipe (2026). Sound-driven practices with digital 3D environments. In S. Gibson, Y. Breuleux, J. Hyde, & D. Leishman (Eds.), *Audio-visualism: Reframing relationships between sound & image* (pp. 337–356). Routledge.

**EDITED BY
STEVE GIBSON, YAN BREULEUX,
JOSEPH HYDE AND DONNA LEISHMAN**



AUDIO-VISUALISM

Reframing Relationships Between Sound and Image

Routledge Advances in Theatre & Performance Studies



Chapter 18:

Sound-Driven Practices With Digital 3d Environments

By **Adriana Sá** ORCID 0000-0003-3919-3121

and **Rui Filipe Antunes** ORCID 0000-0003-3268-3005

Abstract: Digital 3D engines enable the simulation of physical behaviours, conveying immersive cinematic experiences. For example, they can transform the volume and spatialisation of sounds in real-time, in coordination with visual changes or physical actions. In this chapter we look at artistic practices that use these technologies in ways that emphasise the sonic construction. We discuss systems and instruments considering how interaction designs, sound/image dynamics, and audio-visual relationships contribute to shape the perceptual experience. The chapter problematizes the sense of immersion and notion of 3D sound, the superimposition of the physical and the digital space, the dramaturgical potentials of sound spatialization, and the tension between gameplay and musical expression. It overviews a broad range of artistic practices, culminating into a dialogue about the authors' personal quests and approaches.

Quote: Sá, Adriana, & Antunes, Rui Filipe (2026). Sound-driven practices with digital 3D environments. In S. Gibson, Y. Breuleux, J. Hyde, & D. Leishman (Eds.), *Audio-visualism: Reframing relationships between sound & image* (pp. 337–356). Routledge.

An introduction to this chapter

Digital 3D environments, which are widely used in video games, architecture, healthcare and automotive industry, are equally explored in New Interfaces for Musical Expression (NIME).¹ 3D engines enable the simulation of physical behaviours, conveying immersive cinematic experiences. For example, they can transform the volume and spatialisation of sounds in real-time, in coordination with visual changes or physical actions. In this chapter we look at artistic practices that use these technologies in ways that emphasise the sonic construction. We discuss 3D systems and instruments considering how interaction designs, sound/image dynamics, and audio-visual relationships contribute to shape perceptual experience.

Such systems and instruments may use similar technical platforms, but they can be informed by vastly different motivations, challenges and creative processes. In participatory systems, for instance, the audience is central to the design. One important consideration is the limited time users will spend interacting with the system, as Lister et al. explain.² With interfaces conceived for a particular person, who acts as both designer and user, the case is fundamentally different: an iterative process unfolds. For Michael Waisvisz, for example, changing the algorithms from his musical interface meant learning a new instrument³; Joanne Cannon and Stuart Favilla also stressed that creating a new instrument means developing new skills to play it.⁴

As cultural objects, sonic artworks exploring 3D environments entail aesthetic, social and technical layers. To discuss how they can drive a diversity of perceptual experiences, we will examine a set of key-variables: the sense of presence in a mediated environment, the notions of 3D sound and musical motion, space-time composition, the reasons for vision to usually subordinate audition, and the ways through which that sensory dominance can be counteracted.

Armed with this body of knowledge, we will then overview the history of music with 3D environments. We will look at how pioneering practices addressed

networked interaction, the superimposition of digital and physical spaces, and the use of video game paradigms as musical tools. That historical overview will further situate a conversation between the authors: Rui Antunes, whose research focuses on virtual worlds and autonomous agents, will interview Adriana Sá about her transdisciplinary music practice with digital 3D game engines.

Spatial Presence, 3D Sound and Musical Motion

The term ‘spatial presence’, which emerged in media theory to describe the sensation of being physically present in a digital environment, is now equally applicable to any mediated space. This subjective experience reflects how we unconsciously perceive space⁵ and direct our attention^{6 7}, how we become aware of the medium⁸ and suspend disbelief. In short, spatial presence involves interpreting sensory input based on previous experiences, mapping this information to an environment, and adjusting our neural circuits to the focus of our attention.

Sound system designers are highly aware of these processes. The most common sound diffusion system, stereo, uses two speakers or a pair of headphones to produce a 180° auditory image in the horizontal plane. The perceived position of each sound results from the difference between the volume of the left channel and the volume of the right channel, mimicking how we perceive sound sources in the physical world: the ear further from the sound source receives less sound energy due to the head’s shadow. Similarly, Binaural stereo technologies create a 360° auditory image in both the vertical and horizontal planes, by reproducing the time interval between ears, and how the head’s position transforms the spectrum of sound. Furthermore, multiple audio signals can be sent to large speaker arrays. With an Ambisonics⁹ system, these can produce a full-sphere auditory image, covering sound sources on the horizontal plane as well as above and below the listener.

With digital 3D engines, the methods of audio spatialisation are embedded in the digital platform itself, regardless of any output speakers. This makes it

useful to adopt a particular distinction between 2D and 3D sound, as has been proposed by the creators of *FMOD Studio*, a popular 3D positional audio effects software.¹⁰ TV and CD recordings are always 2D: multiple soundtracks may be sent to multiple speakers, but we hear what has been recorded beforehand. That is not the case with 3D engines, where the rendering of sounds and camera views echoes the real-time navigation through a digital environment.

Each digital sound source is placed at a specific location in the digital 3D world, surrounded by an area where sound propagates and fades. When the digital camera view - the image on the screen - is set to equal our listening perspective, the sonic output reflects not only the spatial relation between the sound emitters, but also their position relative to the 3D camera. As the camera or sound source moves, the audio adjusts accordingly by altering how the signal is routed to each output speaker.

The guiding paradigm in the rapid development of new 3D technologies has been to maximise the possibility of creating realistic emulations of our natural relation with the physical world. And then, of course, an artwork might twist that - the variables that influence our subjective feeling of presence can be recombined in unrealistic ways, as happens in soundscape composition or film for example.¹¹ In film theory, Michel Chion coined the term “superfield” to designate how the sound extends the on-screen environment beyond the image frame.¹² Spatial audio does this in a particularly convincing manner, but in fact, a few perceptual cues suffice to produce a sense of immersion in the audio-visual continuum: one does not need a multi-channel sound system to create an auditory image, and one does not need to see a sound source on the screen to infer its location.

Again, the feeling of presence depends primarily on human perception mechanisms; it is not tied to any particular technology. Recognisable sounds, for example, can lead us to feel present in imaginary places, acquiring a symbolic dimension through resemblance, existential connection or causality¹³ : the sound of sea waves evokes the sea, while the sound of an urban environment evokes a city. And the feeling of presence can also emerge when sounds appear unrecognizable, as every sound possesses acoustic properties that carry

information about location and environment.¹⁴ Moreover, the term ‘immersion’ seems applicable to a diversity of musical idioms, from the early massive guitars of Jesus and Mary Chain, to Phill Niblock’s musical drone performances: whilst the overwhelming sound traps your senses, the subtle shifts in texture sharpen your focus.

Sound can be a vehicle to pursue causes and meanings, but we might equally be driven to focus on the sonic experience itself. As mentioned in Chapter 9, Pierre Schaeffer coined the term ‘reduced listening’ to describe this particular mode of perception.¹⁵ While causes, concepts and meanings are important in many musical approaches, they are just one aspect of the experience. Sometimes, they emerge in conscious awareness; other times they fade, as attention focuses on the intrinsic qualities of sound.

These movements of attention are evident when music oscillates between points of low and high intensity. At times it complies with unconscious expectations, producing a sense of continuity; those events are of low intensity, leaving attention under individual control. At other times, unexpected changes monopolise attention, increasing neural activity: those are points of high intensity. A single, recognizable sound in an abstract soundscape creates a semantic discontinuity. But its combination with other events might lead attention to shift somewhere else. For example, when recognizable sounds merge into a musical phrase, attention might drift away from meaning, and focus on timbre, tonality and texture instead. Also, sonic atmospheres can richly intertwine multiple layers of continuity and discontinuity, where subtle pitch variations, varying time intervals and reverberation shifts invite attention to discern musical motion within a seemingly casual mass of sounds.

Musical motion can be forged in many different manners, and music with 3D environments certainly involves vision and space, beyond audition. Simon Emerson uses the term ‘performative arena’ to describe the relation between performer, audience, space, sound sources and events.¹⁶ This term is quite useful to our present discussion; its meaning only needs to be extended so as to embrace images as well. We can say that the performative arena is reflected in the audience’s subjective sense of presence. It can contract and expand, inextricably related to

attentional processes: sometimes attention is drawn to the performer; at other times it shifts to the physical space; and other times the mind flows to imaginary places. Musical motion embeds spatialisation, and spatialisation embeds musical motion.

The Performative Arena and the Grey Areas of Sensory Dominance

Digital 3D systems combine human actors and machinic actors. The listener acts in the world which also acts on itself and affects the listener. This aligns with Alexander Galloway's notion of an *action-based medium*: "...One whose very materiality moves and restructures itself -pixels turning on and off, bits shifting in hardware registers, disks spinning up and spinning down."¹⁷ By adjusting and reshaping the environment, human-machine interactions continually reorganize the listener's experiences and the meanings attributed to them. The placement of sonic objects in space, time and memory can create tension, contrast and conflict.

To some extent, the principles of composing with space and movement in the physical world also apply to digital 3D environments. Whilst the spatial arrangement of sounds and their relative intensity varies depending on how a person moves through a 3D environment, the ways in which the performative arena might contract and expand are a matter of compositional choice.

For example, a participatory work creates a context for the audience to explore. Max Neuhaus puts this in eloquent manner: "I am interested in [...] locating [the sound elements of a composition] in space and letting the listener place them in his own time."¹⁸ In Nye Perry's physical installation named *Exploded Sound* ¹⁹, people move in-between 60 small speakers, experiencing their own, unique mix of sounds and emerging events. Also, audio-visual composition can draw from interactive narratives normally associated with video games; Marie-Laure Ryan identified over a dozen ways through which event patterns can shape tension and release.²⁰ Such participatory works invite people to creatively explore an organised environment, with specific points of interest and paths. Alternatively, when an author performs with their idiosyncratic interface, their timings frame the

audience's experience by shaping a dramaturgical arc within a timeline - as happens in any concert.

Creative processes entail many layers, particularly when instrument, composition and performance are considered as a whole. Several analytical models blur their boundaries^{21 22}, which is particularly useful when we speak of NIME. A typical NIME setup often comprise an input device, mapping algorithms, live input modulation, a structural layer defining musical sections, output channel configuration and digital-to-analogue converters.²³ Acoustical instruments can also become data to be “sensed” and analysed; mapping techniques may turn the microphone into an interactive sensor. And for artists working with digital 3D environments, the moving images, loudspeakers and physical setup are integral to the instrument-system, affecting the overall experience.

A 3D engine is both a component of the system and a system itself. Software mediates our action through code, and code embeds theories that are necessarily biased according to specific criteria. 3D technologies are primarily designed to convey a sense of verisimilitude, and that is inscribed in world-making tools: shaders for real-time visual realism, physics engines for handling collision and body dynamics, positional audio and audio occlusion. Maximising verisimilitude implies maximising the fit between movements, visual components and sonic events. A common approach is to design one-to-one or one-to-many relationships, reflecting expected physical world behaviours.²⁴

These audio-visual correspondences facilitate interaction in participatory systems, help blend the physical and the digital spaces, and enable connections between remotely interacting participants. Consistent interface behaviours make navigation and interaction seem intuitive, effortless and natural. In this way, the material interface is overlooked. As Alexander Galloway notes: the more an interface erases traces of its functioning, the more it succeeds, ultimately becoming as naturalised as air or as common as dirt.²⁵

As a corollary, “dematerialising” the medium might not be a desirable strategy in musical instruments. For example, Tom Mudd et al. speak of “material-oriented approaches”, which focus on nonlinear behaviours and specific sonic properties.²⁶ Volatile, complex interface behaviours engage musicians with the

instrument's materiality, allowing for a bi-directional relationship that wakes latent abilities.²⁷ NIME artists often find that the interaction should be learnable and repeatable to some extent, but complexity and volatility should exceed the performers' absolute control.²⁸

Another important aspect to consider in audio-visual works is how perception handles the combination of auditory and visual stimuli. The artistic debate on visual dominance has proliferated since the 1960s. Thirty years later, Jeff Pressing observed that in digital 3D environments attention prioritises vision whenever causation is discernible, even if the actual sound precedes the image in a new cinematic scene.²⁹ Also, 3D animators often place a footstep sound slightly before the foot actually touches the ground. And research in perception science confirms that a highly congruent audio-visual relationship leads attention to ignore diverging sensory information, while subordinating audition to vision.^{30 31 32} This conveys the primary aim of the brain, which is to detect, perceive and react to events in efficient ways.

Perceptual binding is governed by subconscious concepts of causation. Inconclusive concepts do not necessarily break continuity, because the brain seeks to integrate the heard and the seen into a unified perceptual representation. For example, hearing a glass breaking behind us can evoke a window, a mirror or a dish; and hearing a train while seeing a forest leads us to imagine an environment that includes both. Furthermore, using abstract sounds and images, we demonstrated what we call the *fungible* relationship.³³ It suggests causation because several components convey perceptual binding. Yet, since the cause-effect relationships are unclear, the brain doesn't distinguish fitting and non-fitting information; attention embraces both with inconclusive concepts.

The audio-visual relationship plays a crucial role in sensory dominance, and so do the visual dynamics in themselves. Sudden visual discontinuities can disrupt the experience of complex sonic constructions³⁴: they prompt automatic attention, monopolizing conscious awareness. And the threshold between continuity and discontinuity is fuzzy. For example, an initially disruptive event can create a sense of continuity through consistent repetition, and prolonged exposure can make

small changes in detail feel more, increasing our susceptibility to automatic attention.

The configuration of instrument-systems influences perceptual dynamics, by enabling a field of possible interactions and output results. The range of possibilities can be constrained to greater or lesser extent, based on compositional criteria.³⁵ . The process can be compared to writing a score with a personal music notation system, which an artist can use to prescribe some aspects and leave other aspects open to real-time decision.³⁶

Remote networked interaction and the superimposition of physical & digital spaces

Artists, musicians and programmers have been using digital 3D environments since the late 1990s. Improvement in the internet bandwidths led to an unprecedented expansion of the performative arena, through remote networked interaction. The high-quality digital audio exchange between remote locations was now possible, despite latency.

Vibeke Sorensen, Miller Puckette and Rand Steiger created a system for the *Global Visual Music Project* (1997)³⁷, which combined digital video, 3D graphics, acoustic musical instruments and computer music. This system was played by a chosen set of artists. It emphasised clear connections between sounds and images, linking remote performers, while avoiding direct metaphors like those between pitches and colours ³⁸.

In turn, John Klima's *Glasbead* (1999) was a multi-user system built with a videogame engine.³⁹ The digital interface featured a rotating spherical structure with hammers, stems and bells (Fig. 18.1). Users could import sound files into the bells and trigger them by flinging the hammers. The system allowed for up to twenty online players jamming together remotely. Christiane Paul described it as both 'instrument and toy'; an artwork dealing with the then emerging technologies and practices - multi-user environments, gaming and file-sharing.⁴⁰ The interface

was user-friendly, but no individual user was in command of the sonic construction, which could often incorporate chaotic behaviours.

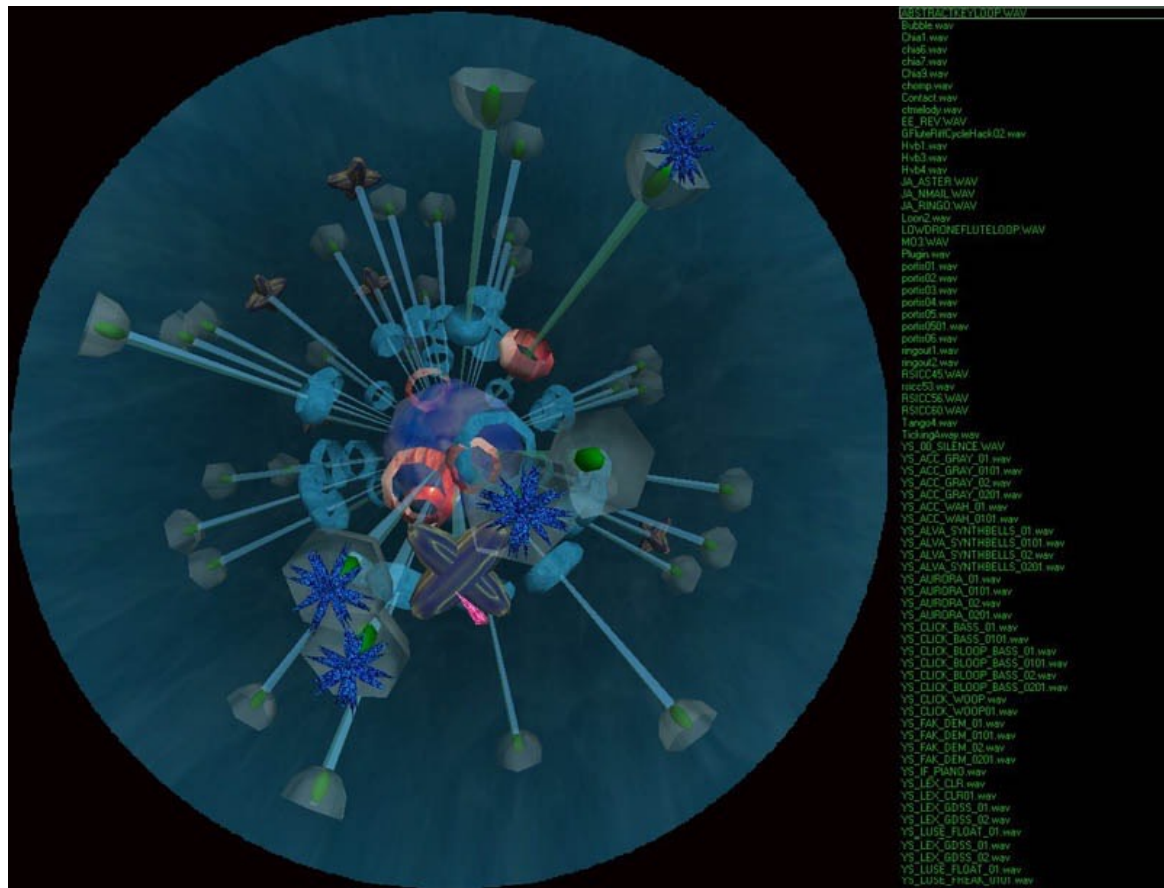


Figure 18.1: John Klima's *Glasbead*, 1999. In <http://www.cityarts.com/glasbead>.

In the works we just described the performative arena ties multiple physical locations, a digital 3D environment and the space of the internet. Other creative approaches focus on superimposing the physical and the digital spaces in a local manner. Char Davies' *Osmose* (1995), a pioneering virtual-reality installation, invited the audience to navigate a digital space using a head-mounted display and motion tracking based on breathing and balance.⁴¹ They could freely explore the environment and control their contemplative experience by moving from one point to another. The audio-visual relationships were fungible and ambiguous. Avoiding the realism typically associated with 3D technologies, Davies explored semi-

abstract, translucent images and flowing particles with the soundscape changing in direct relation with people's movement.

Another, recent, example of spatial superimposition with head-mounted display and motion tracking is Brian Baumbush's *Polytempo Music* (2023) ⁴², a VR application that explores a 48-minute composition. Each instrumental component moves around, leaving a visual imprint of abstract graphics in a darkened 3D space. The sound and the image feel connected, but the basic cause-effect relationships are not straightforward - again, a fungible audio-visual relationship. As listeners navigate through the musical landscape, their position relative to the sound sources defines what they hear. Thus, the listeners' active participation becomes pivotal to their sonic experience.

Other authors are rather driven to refine the capabilities and affordances of a 3D instrument specifically made for musicians. In 2006 Mike Wozniowski, Zack Settler and Jeremy Cooperstock presented *Audioscape*, a system that modelled the user's body in a digital 3D world.⁴³ Controlled by physical activities like moving, turning, pointing or grabbing, it employed physical models of sound propagation to process sound at various locations in the 3D world, but the authors bended the rules of physics for musical purposes.

In the same year, Martin Näf and Daniel Collicott introduced a VR system that sought to address issues surrounding live music performance while adapting software from the automotive industry. ⁴⁴ The system allowed for multiple persons to dynamically position and manipulate digital audio objects in a 3D space, through an interface that included tracking data gloves and a stereoscopic projection displaying the 3D sound "stage".

Fuzzy thresholds between gaming and music

Drawing from the video game culture, the first-person-shooter paradigm has also been adopted as a musical method. In 2007 Mick Grierson introduced a '3D first-person composition and improvisation system' where users could create, adapt and combine elements with varying physical attributes to produce musical structures.⁴⁵

The system entailed no pre-set constraints on the environment, object properties or interactions; it used adaptive algorithms to integrate user input.

The following year, Robert Hamilton presented a system⁴⁶ where weapon projectiles triggered sounds upon collisions. The sonic output reflected kinetic behaviours from physical sound propagation. The system drew from the topography of the virtual environment so that the interactor retained a level of control over the music.

The integration of gameplay mechanics in music creation brings in a delicate balance between control and ‘indeterminacy’. Tarik Barri’s *Versum*⁴⁷ (developed since 2008) also explores first-person navigation and camera view, but the dynamics of electronic sounds and abstract graphics are unrelated with “shooting”. The sound emitters are displayed as colour beams fluctuating in space (Fig. 18.2); each sound is audible within an invisible area around a visible beam. Technically, the sound spatialisation reflects which beam produces each sound. Nevertheless, there is some ambiguity because invisible sound areas intersect with each other and persist after their visual source disappears from the screen.



Figure 18.2: Tarik Barri’s *Versum*, 2008. In <http://tarikbarri.nl/projects/versum> .

Repurposing videogame technologies for music composition and performance also raises questions about the notion of ‘play’ in gaming and music. Florent Berthaut et al. used the first-person-shooter paradigm in *Couacs* (Fig. 18.3),⁴⁸ a ‘multiprocess instrument’ for collaborative musical interaction introduced in 2011. It led them to ask:

How can we use some game actions for musical control without disturbing other game actions not connected to sound, and vice-versa? Will gamers/musicians try to learn how the instrument works and how they can produce specific musical results, or will they only play without paying attention to the generated music? Should these instruments have a goal like a video game or not?⁴⁹



Figure 18.3: *Couacs*, by Florent Berthaut et al., 2011. In https://www.researchgate.net/figure/Three-musicians-playing-the-Couacs-in-Faders-For-All-mode_fig1_228985639 . Used by permission

Rather than using general-purpose interfaces like keyboard and mouse, Christos Michalakos (since 2016) explores an augmented, electro-acoustic drum-kit, which provides sonic content while also working as a game controller. According to him, this musical interface offers a much wider range of expressive possibilities, control and tactile feedback than general-purpose game controllers. The documentation video⁵⁰ shows how the actual cause-effect relationships remain obscure to the audience - another fungible relationship.

Game technologies are well integrated in the musical world at present. But something unsettled lingers between gaming and music, a tension that never resolves. Psychologist M. Csikszentmihalyi elaborated a theory of flow ⁵¹ that became widely influential in game design.⁵² Some aspects of that theory are applicable in the musical domain, while others seem arguable.⁵³ Indeed, practitioners differ on whether the notions of ‘play’ are equivalent in games and music. One might be inclined to dissolve the difference, or rather sense a conflict. In both music and gaming, flow involves concentration, action awareness, and transformation of time; but whereas a game player needs clear purposes, a musician focuses on perceptual experience itself. Whilst a game player loses self-consciousness, a performer remains highly aware of how their presence informs the audience’s experience.

The tension between gaming and music through the lens of a personal practice

The terms one uses to discuss creative works can resonate with different concepts, experiences and implications. As such, this chapter articulates two interpretative lenses - Rui Antunes’ stems from evolutionary art and virtual worlds, while Adriana Sá’s stems from transdisciplinary music. Convergences and divergences between their standpoints - which seemed quite fuzzy at first - became clearer once they started discussing her artistic practice.

Rui Antunes - You created an audio-visual instrument that combines an ancient string instrument, the zither, and 3D software, which processes sound and graphics based on the zither input. This involved the repurposing of technologies intended for the creation of video games. Could you explain the ideological and political dimension you see in circumventing the purposes digital tools are designed for?

Adriana Sá: Usually we handle computers as magic black boxes, but software mediates our action through code. Code embeds theories that are necessarily biased according to specific criteria, which are too often taken for granted. The more science and technology are involved, the opaquer and obscure they become, and the more distant we become from computation as creative material - especially if we are not coders. Circumventing the theories governing digital tools is a way to emancipate our interaction with the world.

RA - Can you expand on how this idea manifests in your artistic work?

AS: An instrument indicates a hierarchy of concerns, which determine its functions and meanings. These might be explicit, implicit, or even hidden. With novel instruments, inexplicitness and equivocalness are as important as in any artwork.

RA: So, you are driven to obscure causal relationships. And yet, simultaneously you want the audience to be engaged in understanding the system. Is this part of your aesthetic system? Are you proposing an experience that is almost gamified in the sense that part of the enjoyment comes from deciphering how the circuit works?

AS: No, not really. I expect people to quit trying to understand the mechanics of the instrument and focus on perceptual dynamics itself. The relation between sound and image conveys perceptual binding, but it also loosens the perceptual hierarchy so that attention embraces fitting and non-fitting information - with inconclusive concepts.

RA: Jay Bolter and Richard Grusin proposed the term 'remediation' to describe the representation of one medium in another, arguing that contemporary media acquires cultural significance while revising, repurposing,

remixing, referencing and/or competing with older media.⁵⁴ Perhaps we can equally speak of ‘remediation’ when an artist bridges their ‘old’ and ‘new’ mediums - additional levels of significance might emerge. How does your audio-visual instrument relate to your previous work, which is strongly site-specific? Is there a continuum?

AS: In the late 1990s I started using sensor technologies and software for connecting sound with light, space, movement, architecture, weather and social context. I frequently used sounds collected onsite. My current instrument also processes pre-recorded sounds - a way to expand the performative arena beyond the physical performance space. In my previous work, I also developed the “bodiless instrument”: a set of amplified wires stretched in space, attached to the walls; it acquired different timbral qualities in every performance-installation. As I moved in space to play the wires, my shadow upon the light sensors also affected the processing of digital sounds. Neither the digital nor the acoustic components of my setup were fully controllable. I found this stimulating in performance, because the interaction with the system felt reciprocal, and I needed a particular state of concentration to create musical meaning upon the unexpected. That is a crucial creative motivation in my subsequent work as well.

RA: It seems like you're embracing indeterminacy. Are you intentionally loosening the direction and control of the piece? Is this approach part of your poetics?

AS: Whilst the zither functions as a controlled prolongation of the body, unexpected digital sounds expand my actions beyond conscious intention. An unexpected, often minute event can produce compelling performative tension. It causes a minimal, yet graspable hesitation – a moment of suspense. Resolving the musical challenge in good time then causes a sensation of release. To address the unexpected resourcefully, performative action must exceed intellectual deliberation - musical timings and intervals possess biophysical logics.

RA: Video game technologies embed theories that seem in contradiction with your ideas about interface behaviours and perceived cause-effect relationships. How did you reconcile those theories with your creative concerns?

AS: When I began to work with video game engines in 2007, I had to question certain theories embedded in the digital platform. The software assumed that digital behaviours must be consistent and cause-effect relationships clearly perceivable. My creative concerns are very different, and I wanted to extract musical meaning from that divergence. Additionally, I wanted the image to extend the physical world in depth without distracting attention from the sonic construction. And that seemed problematic because I had the strong impression that vision tends to subordinate audition.

Drawing from audio-visual theory, psychology, neuroscience, interaction design and musicology, I found three creative principles: a) to calibrate thresholds between human control and machinic unpredictability, so as to create sonic complexity; b) to dispense with disruptive visual changes, so as to prevent the dynamics of the image from monopolising attention; and c) to create a fungible audio-visual relationship, where attention embraces fitting and non-fitting information because the cause and effect mechanics are unclear.

Throughout its multiple iterations, my instrument has been exploring these principles in different ways.⁵⁵ Currently I am particularly interested in grey areas, such as the fuzzy thresholds between continuity and discontinuity.

RA: Your collaborator and husband John Klima was amongst the first artists using video game technologies in interactive systems and installations. Did this inspire the emergence of your instrument? When you met in 2007, your artistic paths seemed quite different. So how did they intersect?

AS: Klima inspired me to combine an acoustic instrument with audio-visual 3D software, as he created software that processes graphics and recorded sounds based on audio input, through amplitude and frequency analysis. That software was ironically titled *Portuguese Guitar Hero*, after the popular Guitar Hero video game series. Originally, Klima was driven to explore the gamification of music and the audience's participatory interaction - as in his *Glasbead*, from 1999. I wanted to take advantage of the technical connection between analogue and digital components, but needed to transform digital behaviours and aesthetics according to my own creative concerns; and Klima agreed to change the code accordingly.

The software became tied to a specific, custom zither input. The zither has aged strings and custom (de)tuning, allowing for unpredictable digital reactions.

RA: What specific theories embedded in video game engines did you find most limiting or challenging when you started developing your instrument?

AS: Klima and I began our practical investigations with an open-source game engine called Torque, coupled with the FMOD audio library. The idea was to use two interfaces: the audio analysis of the zither and a set of gamepad switches. We distributed sound emitters throughout the terrain of a 3D world. Navigating through the digital space would make them audible, and I enjoyed how sound spatialisation was bound to their visual position on the screen. However, dwelling with musical timings was cumbersome because navigating the digital space from one point to another takes time, and because the interaction with the system required too much visual attention; I wanted to focus on the sonic construction without needing to look at the screen.

RA: Did this process evolve? Was there any iterative process and learnings?

AS: The sound emitters were then mapped to particles, and that worked better because I could forge dynamic visual behaviours without navigating through the digital space (Fig. 18.4). The 3D camera/ listener was configured to rotate in a single spot. That enabled me to prevent unwanted changes in the volume of digital sounds, while retaining a connection between the visual dynamics and the dynamics of sound spatialisation.

RA: To which extent did technical solutions stem from your zither playing techniques? Did the tension analog-digital lead to any creative discoveries?

AS: The first version of my audio-visual instrument used the Torque game engine, which allowed me to change the aspect of the 3D world and map my sounds to existing visual events. Yet, the theories governing the interaction design assumed that music deals with tones and halftones, ignoring the chromatic wealth that exists between them. I like to explore those in-between territories: my zither has aged strings and a peculiar tuning, and I play it hands, cello bow, pick or bottlenecks. The solution was in the fact that digital behaviours were contingent

to amplitude and frequency analysis, which means that the zither input could lead the software to behave in ‘chaotic’, non-linear ways. In turn, the relation between acoustic and digital sounds led me to discover a new musical idiom, which intertwines ambivalent timbres in orchestral ways.



Figure 18.4: Two performances from the *Windowmatter* series, which explored an early version of the audio-visual instrument, with the AG#1 software. Top - a solo by Adriana Sá at In-Sonora, Madrid, 2010. Bottom - A. Sá and John Klima with a networked version of the software, *15th anniversary of Ernesto de Sousa Fellowship*, Lisbon, 2008. Used by Permission.

RA: What is the role of the image?

AS: The projected image is a shifting camera view over a digital 3D world, which morphs upon audio input detection. In performance I stand in front of the screen, facing the audience and wearing white clothes. The image is desired to create a reactive stage scene, and this requires the audio-visual relationship to produce a sense of causation. Simultaneously, the audience is desired to focus on the sonic construction, which requires the concepts of causation to remain inconclusive.

RA: This relates to the material determinism you mentioned earlier: you are obviously open to discover, and counteract, the properties of any given medium. You clearly do not wish to sacrifice your sensibility, but you use technologies governed by very different ideologies. Is there also a level of commitment with algorithmic and machinic behaviours?

AS: The Torque engine (now discontinued) entailed several functions that were irrelevant to my instrument, and I didn't have the coding skills to get rid of them - a waste of computer memory, so precious in real-time rendering. Yes, I transformed this obstacle into creative material, but the result was not fully satisfactory.

The new audio software used several layers of digital analysis to produce sonic complexity. The mathematical difference between the detected pitch and the closest tone/half tone turned out to have an enormous musical potential: that value could be applied to a pitch down effect. In this way, I could mingle pre-recorded sounds with inter-tonal transformations and create new musical forms. This led me to develop specific combinations of zither techniques and digital sounds, interlacing, sometimes merging, other times diverging - like a braid. This is an archetypical form, which cannot be reduced to any musical style.

RA: Working at low-level in the digital architecture led to new sonic constructions. Did it equally influence the complexity of the audio-visual mapping, given your concerns with visual dynamics and sonic experience?

AS: Yes, it did: the digital 3D world became much more organic, and so did the audio-visual relationship. Once the audio architecture was established, the

digital prototype was implemented within a 3D engine written from scratch. Now the sound emitters emerged as terrain elevations, while the 3D camera/ listener rotated in a single spot. Multiple layers of digital analysis were applied to the visual dynamics, creating a wealth of ambivalent discontinuities. By overlapping multiple images with transparencies, I could accentuate the impression of an organic, moving painting (see for example Fig. 18.5).

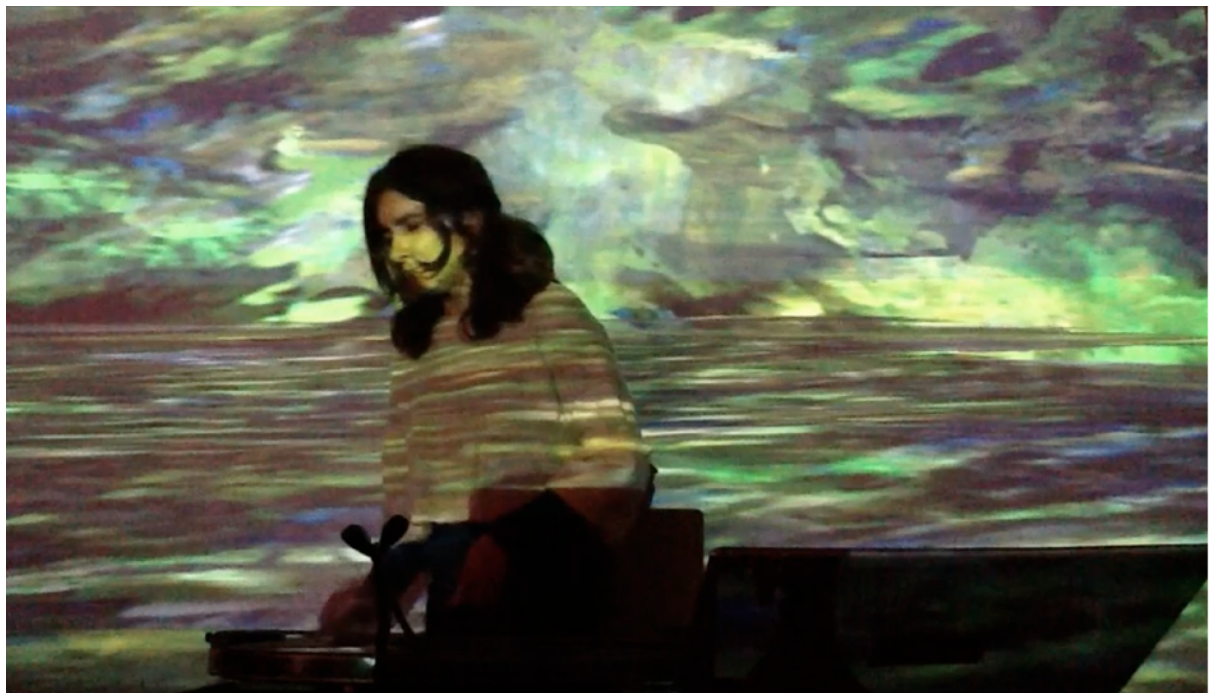


Figure 18.5: Adriana Sá performing with the AG#2 software at STEIM during the Musical Organics Symposium, Amsterdam, 2016. Used by Permission.

Recent versions of this audio-visual software use the same code and the same parameter configurations. They make use of yet another 3D engine, called Irlicht, combined with the FMOD audio library.

The core mechanics of the instrument are now crystalised, but the instrument continues to evolve as an open system: new recorded sounds and new zither playing techniques give rise to new musical forms, and new images applied to the 3D world give rise to new stage scenes (see for example Fig. 18.6).⁵⁶

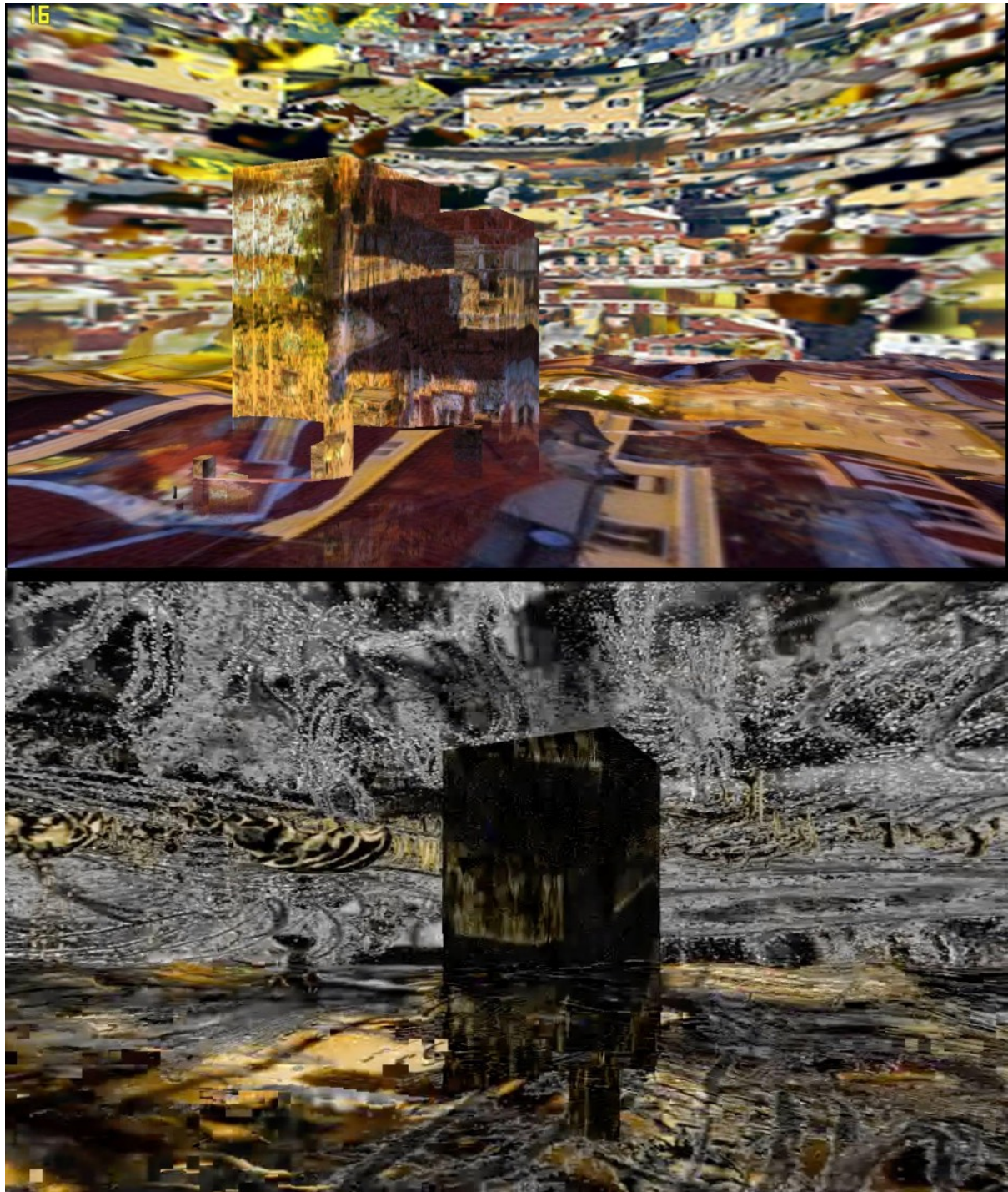


Figure 18.6: Top - screenshot from AG#3.3 (2020). Bottom - screenshot from AG#3.9 (2022). Both software versions are made with a 3D engine called Irrlicht and an audio library called FMOD. Used by Permission

RA: Now that the design of your instrument seems stabilized, how do you balance the need for visual interaction with your desire to focus on sonic construction without looking at the screen?

AS: The zither is audible, and the microphone attached to it works as a sensor - the sound becomes data, which can be mapped to any digital events. My instrument explores disparities between the direct output of the zither, and the indirect outputs of digital sound and image. The audio-visual mappings are prepared prior to performance; digital constraints assure visual continuity without my needing to look at the screen.

Concluding...

Through a discussion centred on Sá's audio-visual instrument, we explored the tension between gaming and music, expanding on the interplay of concepts introduced in the previous sections of the chapter. By repurposing game technologies, Sá disrupts the deterministic logic embedded in digital tools, fostering an ambiguous relationship between sound and image that encourages perceptual engagement rather than structured interpretation.

Beyond this particular creative position, there are certainly a great diversity of artistic lenses, motivations and approaches. The question then is what makes all these practices genuinely new. By looking at sound spatialisation, human-computer interaction, media theory and perception sciences, we sought to identify conceptual tools that help grasp the “newness” in operative modes, and the dynamics at play in this liminal, interstitial field. We further overviewed the history of music with 3D environments, while looking at a wealth of artistic approaches. Taken together, these explorations suggest that the relevance of contemporary digital 3D performance lies not in any single technology but in hybrid, cross-disciplinary strategies that interweave spatialized sound, interactive systems and perceptual engagement — shaping the relationships between performer, audience and environment while probing the boundaries between the physical and digital worlds.

Endnotes

- ¹ Official website: <https://nime.org> (accessed September 2025)
- ² Lister, M., Dovey, J., Giddings, S., Grant, I. & Kelly, K. (2009) *New Media: A Critical Introduction* (2nd Edition). Routledge, pp. 22-24. Available at: <https://www.routledge.com/New-Media-A-Critical-Introduction/Lister-Dovey-Giddings-Grant-Kelly/p/book/9780415431613?srsltid=AfmBOopQb-OGii7QL0YX0iUWJudagLV5ejzTY8R4pD51FubDCZ7cF3c9>
- ³ Waisvisz, M (1999) "Gestural Round Table" in *STEIM Writings*.
- ⁴ Cannon, J. & Favilla, S. (2012) "The Investment of Play: Expression and Affordances in Digital Musical Instrument Design," in *Proceedings of ICMC - International Computer Music Conference*, Ljubljana, pp. 459-466. Available at: <https://quod.lib.umich.edu/i/icmc/bbp2372.2012.086/1>
- ⁵ Schubert, T.W. (2009) "A New Conception of Spatial Presence: Once Again, with Feeling," *Communication Theory*, 19(2), pp. 161-187.
- ⁶ Draper, J.V., Kaber, D.B. and Usher, J.M. (1998) "Telepresence," in *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(3), pp. 354-375.
- ⁷ Sá, A., Tanaka, A. (2023) "A Parametric Model for Audio-Visual instrument Design, Composition and Performance," in *Live visuals: History, Theory, Practice*. Edited by D. Leishman, S. Arisona, and S. Gibson. Abingdon, Oxon: Routledge, pp. 240-266.
- ⁸ Hartmann, T. (forthcoming, 2026) "Being Present in Virtual Reality and Augmented Reality: Explicating the Psychology of Immersive Technology (XR)," in *De Gruyter Handbook of Media Psychology*. Edited by Bailey, R. & Glenna L. Read. Berlin: Germany. Preprint available at: https://www.researchgate.net/publication/381776456_Being_Present_in_Virtual_Reality_and_Augmented_Reality_Explicating_the_Psychology_of_Immersive_Technology_XR (accessed May 2025).
- ⁹ Gerzon, M. (1985) "Ambisonics in Multichannel Broadcasting and Video," in *Journal of the Audio Engineering Society*, 33(11), pp. 659-871.
- ¹⁰ "3D Audio" (2014) in *FMOD Studio 101*, p. 296. Available at: http://www.soundlibrarian.com/uploads/3/1/0/6/3106267/fmod_studio_101_chapter_10_3d_audio_gdc.pdf (accessed May 2025).
- ¹¹ For a systematic analysis of how a creative work - sonic, visual or audio-visual - might handle the variables that influence spatial presence, see the method proposed in Sá and Tanaka (2022).
- ¹² Chion, M. (1994) *Audio-Vision: Sound on Screen*, New York: Columbia University Press, p.150.
- ¹³ Mitchell, W.J.T. (2005) "There Are No Visual Media," in *Journal of Visual Culture*, 4(2), pp. 257-266.
- ¹⁴ Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. The MIT Press.
- ¹⁵ Schaeffer, P. (2002) *Traité des Objets Musicaux: Essai interdisciplines*. Paris: Edition du Seuil. First edition published by Editions du Seuil, 1966.
- ¹⁶ Emerson, S. (2017) *Living Electronic Music*. London: Taylor and Francis, p. 92.

- ¹⁷ Gallaway, A. (2006) *Gaming: Essays on Algorithmic Culture*, University of Minnesota Press, p.3.
- ¹⁸ Cox, C. quoting Neuhaus (2006) in "From Music to Sound: Being as Time in the Sonic Arts." *Sound: Documents of Contemporary Art*. Edited by Colab Kelly. Whitechapel Gallery/MIT Press, 2011, pp. 80-87..
- ¹⁹ Parry, N. (2012) *Exploded Sound*. Documentation available at: <https://www.nyeparry.com/portfolio/exploded> (Accessed April 2025).
- ²⁰ Marie-Laure R. (2001) *Narrative as Virtual Reality*. The John Hopkins University Press, pp. 246-264.
- ²¹ Birnbaum, D., Fiebrink, R. Malloch, J. & Wanderley, M. (2005) "Towards a Dimension Space for Musical Devices," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Vancouver, BC, Canada.
- ²² Ciciliani, M. & Mojzys, Z. (2015) "Evaluating a Method for the Analysis of Performance Practices in Electronic Music," in *Inter-Face: International Conference for Live Interfaces 2014*. Edited by A. Sá, M., Carvalhais & A. McLean. Porto University, CECL & CESEM / NOVA University, MITPL/ University of Sussex, pp. 172-182.
- ²³ Tanaka, A. (2009) "Sensor-Based Musical Instruments and Interactive Music," in *Creative and Performance Modes*. Edited by R. Dean. The Oxford Handbook of Computer Music, pp. 233-257.
- ²⁴ See for example *LyraVR and EXA: The Infinite Instrument*, in which the affordances seem to be drawn from the physical world. Available at: https://store.steampowered.com/app/606920/EXA_The_Infinite_Instrument (accessed May 2025).
- ²⁵ Alexander G. (2012) *The Interface Effect*. Polity Press, p.25.
- ²⁶ Mudd, T., Holland, S., Mulholland, P. & Dalton, N. (2015) "Nonlinear Dynamical Systems as Enablers of Exploratory Engagement with Musical Instruments," in *Inter-Face: International Conference for Live Interfaces 2014*. Edited by A. Sá, M., Carvalhais & A. McLean. Porto University, CECL & CESEM / NOVA University, MITPL/ University of Sussex, pp. 172-182..
- ²⁷ Sá, A. (2017) "Designing Musical Expression," in *Proceedings of xCoAx*.
- ²⁸ Sá, A., Ryan, J., Van Der Heide, E., Tanaka, A., McPherson, A., Magnusson, T., McLean, A., Carvalhais, M. & Gierson, M. (2015) "Live Interfaces: Seeds of Debate," in *Inter-Face: International Conference for Live Interfaces 2014*. Edited by A. Sá, M., Carvalhais & A. McLean. Porto University, CECL & CESEM/ NOVA University, MITPL/ University of Sussex, pp. 13-29.
- ²⁹ Pressing, J. (1997) "Some Perspectives on Performed Sound and Music in Virtual Environments," in *Presence*, 6(4), p.8.
- ³⁰ Boltz, M. (2004) "The Cognitive Processing of Film and Musical Soundtracks," in *Memory & Cognition* 32, pp. 1194-1205.
- ³¹ Spence, C. (2009) "Explaining the Colavita Effect", in *Progress in Brain Research* 176. Edited by S. Srinivasan Elsevier, pp. 245-58.
- ³² Robinson, C., Barnhart, W. & Rivera, S. (2015) "Auditory Stimuli Slow Down Responses and First Fixations: Support for Auditory Dominance in Adults," in *37th Annual Meeting of the Cognitive Science Society*, Pasadena, California.

- ³³ Sá, A., Caramieux, B. & Tanaka, A. (2015) "The Fungible Audio-Visual Mapping and its Experience," in *Journal Of Science And Technology Of The Arts* (6)1, pp. 85-96.
- ³⁴ Sá, A. (2013) "How An Audio-Visual Instrument Can Foster the Sonic Experience," in *Live Visuals for Performance, Gaming, Installation and Electronic Environments / Leonardo Electronic Almanac* (19)3. Edited by L. Aceti, S. Gibson, S. Arisona & O. Sahin, O. MIT Press, pp. 284–305.
- ³⁵ See Magnusson, T. (2010) "Designing Constraints: Composing and Performing with Digital Musical Systems," in *Computer Music Journal* 34(4), pp. 62–73.
- ³⁶ More about this idea can be found in Sá, A. (2024) "Transdisciplinarity, Composition, Expression: Reflections of a Spherical Way of Thinking," *Leonardo* 57(1), MIT Press, pp. 33-39.
- ³⁷ Sorensen, V., Puckette, M. & Steiger, R. (1997) *Global Visual Music Project*. Documentation available at: <https://muse.jhu.edu/article/185790> (accessed May 2025).
- ³⁸ The system used Pure Data, the visual programming language developed by Puckette.
- ³⁹ Klima, J. (1999) *Glasbead*. Documentation available at: <https://arteedadsilicio.com/2000-2/2000-john-klima-glasbead> (accessed May 2025).
- ⁴⁰ Paul, C. (2003) *Digital Art*, Thames & Hudson.
- ⁴¹ Davies, C. (1995) *Osmose*. Documentation available at: <https://www.immersence.com/osmose> (accessed May 2025).
- ⁴² Baumbush, B. (2023) *Polytempo Music*. Documentation available at: <https://ccrma.stanford.edu/events/brian-baumbusch-polytempo-music-new-interactive-vr-music-application> (accessed May 2025).
- ⁴³ Wozniowski, M., Settel, Z. & Cooperstock, J. (2006) *Audioscape*. Documentation available at: <http://wikibin.org/articles/audioscape.html> (accessed May 2025).
- ⁴⁴ Näf, M. & Collicott, D. (2006) "A VR Interface for Collaborative 3D Audio Performance," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Paris, pp. 57-60.
- ⁴⁵ Grierson, M. (2007) "Noisescape: A 3d Audiovisual Multi-User Composition Environment," in *The World as Virtual Environment*. Edited by J. Birringer, T. Dumke & K. Nicolai. Dresden: Trans-Media-Akademie Hellerau, pp. 160-168.
- ⁴⁶ Hamilton, R. (2008) "q3osc: Or How I Learned to Stop Worrying and Love the Game," in *Proceedings of the International Computer Music Association Conference*.
- ⁴⁷ Barri, T. (since 2008) *Versum*. Documentation available at: <http://tarikbarri.nl/projects/versum> (accessed May 2025).
- ⁴⁸ Berthaut, F., Katayose, H., Wakama, H. Totani, N. & Sato, Y. (2011) "First Person Shooters as Collaborative Multiprocess Instruments," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 44-47.
- ⁴⁹ Berthaut et al. (2011) p.45.
- ⁵⁰ Michalakos, C. (2016) "Pathfinder: a performance-game for the augmented drum-kit," in *International Conference on Live Interfaces*. Edited by T. Magnusson, C. Kiefer & S. Duffy. Brighton: EMuTe Lab/ University of Sussex & REFRAME 2016, p. 268-269.

⁵¹ Csikszentmihalyi, M. (1996) *Creativity: Flow and the Psychology of Discovery and Invention*. HarperCollins Publishers.

⁵² See, for example, Koster, R. (2004) *Theory of Fun*, Paraglyph Press (excerpts including foreword by Will Wright available at <http://www.theoryoffun.com>) or Sweetser, P. & Wyeth, P. (2005) “GameFlow: a model for evaluating player enjoyment in games,” in *Computers in Entertainment (CIE) - Theoretical and Practical Computer Applications in Entertainment* (3)3 (2005), p. 3.

⁵³ Sá, A. (2014) “Repurposing Video Game Technologies for Musical Expression,” in *Proceedings of the Conference on New Interfaces for Musical Expression*, London.

⁵⁴ Bolter, J.D. & Grusin, R. (2000) *Remediation: Understanding New Media*, MIT Press.

⁵⁵ Early research documentation available at: <https://adrianasa.org/research/practiceOverview.htm> (accessed May 2025)

⁵⁶ See for example Sá, A. et al. (2021-23) *Favela Real, Project Series*. Documentation available at https://adrianasa.org/docu_img/favelareal.htm (accessed May 2025).