



HAL
open science

From the Lab to the Stage: Practical Considerations on Designing Performances with Immersive Virtual Musical Instruments

Victor Zappi, Florent Berthaut, Dario Mazzanti

► To cite this version:

Victor Zappi, Florent Berthaut, Dario Mazzanti. From the Lab to the Stage: Practical Considerations on Designing Performances with Immersive Virtual Musical Instruments. *Sonic Interactions In Virtual Environments*, Springer VS, In press, 978-3-031-04020-7. hal-03790049

HAL Id: hal-03790049

<https://hal.science/hal-03790049v1>

Submitted on 29 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

From the Lab to the Stage: Practical Considerations on Designing Performances with Immersive Virtual Musical Instruments

Victor Zappi, Dario Mazzanti and Florent Berthaut

Abstract Immersive Virtual Musical Instruments (IVMIs) lie at the intersection between Music Technology and Virtual Reality. Being both Digital Musical Instruments (DMIs) and elements of Virtual Environments (VEs), IVMIs have the potential to transport the musician into a world of imagination and unprecedented musical expression. But when the final aim is to perform live on stage, the employment of these technologies is anything but straightforward, for sharing the virtual musical experience with the audience gets quite arduous. In this chapter, we assess in detail the several technical and conceptual challenges linked to the composition of IVMI performances on stage, i.e., their *scenography*, providing a new critical perspective on IVMI performance and design. We first propose a set of dimensions meant to analyse IVMI scenographies, as well as to evaluate their compatibility with different instrument metaphors and performance rationales. Such dimensions are built from the specifics and constraints of DMIs and VEs; they include the level of immersion of musicians and spectators, and provide an insight into the interaction techniques afforded by 3D user interfaces in the context of musical expression. We then analyse a number of existing IVMIs and stage setups, and finally suggest new ones, with the aim to facilitate the design of future immersive performances.

1 Introduction

Making music while immersed in a Virtual Environment (VE) is an exciting experience. In a synthetic space designed to replicate and transcend our world, we gain

Victor Zappi
Northeastern University, Boston, MA, United States, e-mail: v.zappi@northeastern.edu

Dario Mazzanti
Fondazione Istituto Italiano di Tecnologia, Genoa, Italy e-mail: dario.mazzanti@iit.it

Florent Berthaut
University of Lille, France e-mail: florent.berthaut@univ-lille.fr

the ability to become composers and performers of inventive musical pieces, that leverage unprecedented acoustical phenomena (virtual sound), mechanical phenomena (virtual interaction) and perceptual/cognitive phenomena (virtual experience). This is possible thanks to the design of virtual devices that, alike Digital Musical Instruments (DMIs), transform interaction into sound, but also allow to channel musical expression through the peculiar features of the surrounding VE. For example, in a world where sound can be visible as well as tangible (like the one created by Lanier to host his “Virtual Instrumentation” [49]), virtual musical devices might permit not only to play notes but also to manipulate them, as they are still echoing in the air. In other VEs, their design might leverage the possibility to fly or teleport through space to create huge distributed instruments, that would be otherwise impossible to manoeuvre. What these apparently disparate musical applications share is *immersion*, meaning that physical equipment, virtual content and the overarching logic that binds the two are geared towards the amplification of the physical and cognitive involvement of the musician acting in the environment. Being both DMIs and elements of immersive environments, or even extending across whole VEs, we call these devices Immersive Virtual Musical Instruments (IVMIs); they typically consist of virtual representations of sound processes and parameters [11, 35], and rely on immersive multimodal technologies to support fine 3D interaction.

As it happens in the case of most musical instruments, composition or studio practice with IVMIs may be considered an end itself [88]. The musician that is immersed in the VE may experience the feeling of satisfaction typical of the completion of challenging musical tasks [58]. Moreover, in this scenario satisfaction is likely to be combined with the sense of *discovery* that characterises Virtual Reality (VR), as the IVMI may feature novel musical affordances [10] or, more in general, unusual sets of sensory-motor contingencies [79]. While some musical VEs are designed specifically to elicit such autotelic responses¹ [35, 96, 72], the final aim of a considerable number of musicians is to perform with their IVMI in front of various audiences and use it to create some sort of *connection* with them.

Unfortunately, the step from the studio to the stage is anything but straightforward. First of all, the rehearsal spaces of most IVMI players are not standard music studios. Before the release of consumer Head-Mounted Displays (HMDs) and the rise of VR videogames, IVMIs were almost exclusively designed (and played) in VR research facilities, equipped with minimal audio gears like an audio interface and a mixer [56, 10, 94]. Nowadays immersive technology is more affordable and VR musicians may have access to spaces more affine to traditional music facilities [36, 53]; however, in these studios professional audio equipment still needs to be laid side by side with tracking systems, HMDs, projectors, all connected to dedicated computers. The showcasing of live IVMI performances—even the simplest ones—inevitably relies on the employment of such heterogeneous studio equipment, which has to be moved to the venue and arranged on stage.

But it is not the size nor the complexity of the setup alone to qualify IVMI performances as challenging artistic endeavours. Rather, the real hurdle for VR artists

¹ We may label this kind of musical VEs with the umbrella term *installations*.

comes from the nature of the required equipment. Musicians active in contemporary popular and underground music scenes are no stranger to the on-stage employment of remarkably complex setups, the most straightforward example being a generic rock band playing electric, electronic as well as acoustic instruments altogether during the same show. However, when a rock band steps on stage and the concert begins, the chosen technological setup immediately proves itself fundamental to support musical expression and to create a synergetic connection with the audience. The instruments fuse with the bodies of the members of the band and each gesture acquires a clear musical meaning; cables and speakers disappear from sight, concealed by the music and by the flashing lights, that equally immerse the musicians and the audience into a large shared audio/visual environment. Unfortunately, this *inclusive* scenario is in stark contrast with the musical experience that is delivered on average through an IVMI performance. When inside the VE, the musician fully leverages the immersive technology laid on stage to approach the IVMI's logic and control it, as if the instrument were physically there in front of them. But to the eyes of the audience, this interaction is quite cryptic, even abstract. This is due to the fact that spectators are confined outside of the VE. They see the musician assuming awkward poses and contorting themselves while handling invisible objects on a semi-empty stage. The only visible clue about the existence of the VE is the technology that surrounds the performer, which mediates the interaction between the physical and the virtual world, but tells very little about the mechanics of the latter. Without a clear view of the virtual objects and their response to interaction, what remains is just a music piece almost completely disconnected from the gestures and the physical presence of the musician.

Some may argue that the potential of VR music extends way beyond the virtualisation of the sole performer's experience. For a moment we can forget about stages equipped with immersive gears and even venues, and rather imagine shows taking place in completely synthetic worlds, that spectators access remotely from their living rooms. This is one of the many social expressions of contemporary VR culture [62]. Showcasing a fully virtual performance surely helps the audience see the show in its entirety, and better appreciate interaction and aesthetic nuances. However, research proved that VR setups and networked technologies available today are not yet capable of providing the same sense of connection triggered by social activities set in physical reality, let alone by complex psychophysiological experiences like concerts and live performances [50].

Leveraging immersion to create a sense of connection/inclusion clearly becomes the main challenge for VR artists and IVMI designers, and the very immersive equipment they rely upon seems to get in the way. During the first experiments² with IVMIs carried out in the early '90s, this scenario was not necessarily considered a limitation to artistic expression. Conversely, it was taken as an opportunity to explore a new relationship between performer and audience [49]. Immersive technology was employed to create novel musical instrumentation, but at the same time concealed

² Here the use of the term "experiments" does not want by any means to devalue the artistic significance of these early performances; it refers to the inability to predict the effects that such a novel technology would have had on the overall vibe of the shows and on the audiences' experience.

this instrumentation and made the musician unable to know how this looked like to the audience. Differently from a traditional musical performance, this scenario did not elevate the status of the performer; it put performer and audience on the same level instead, providing IVMIs with a distinct aesthetic. However, while a similar peer relationship could be achieved by means of other forms of live art too [49], some emerging properties of VR appealed new media artists for their utter uniqueness [98, 57]. This caused a rapid change of paradigm and by the mid '90s the focus of VR performances became revealing a new world, as opposed to concealing it. Immersive technology started to be praised for its potential to offer to the audience a stake in the VE, in the form of a *vicarious* experience [31]. In other words, for the first time VEs were conceived as spaces where to live an experience not only via first-person interaction (as in the case of the performer), but also by observing someone else interacting. This emphasised the importance of being able to perceive continuity between the performer's gestures and the resulting sounds and visuals, to connect with the performer, and share with them the same virtual musical experience.

Today, the design of most IVMIs and VR performances is based on the same rationale [10, 36, 41], as artists aim for shared experiences and connection with the audience. But in practice this is a costly goal. The setting up of such musical VEs requires beforehand a strong commitment to understanding both what performing music means and how VR affects action and cognition, as well as a fair amount of equipment ready at hand. Alas, in real-case scenarios mental and physical resources tend to be limited; trade-offs happen to be extremely common, either in terms of a reduction in the sense of agency and immersion of the performer (to favor the audience's side), or as an overall depreciation of what attending *live* music truly feels like (biasing the performer's role).

1.1 The Role of Scenography

The impending gap between performer's and audience's virtual musical experiences is a complex phenomenon that has to be accounted for in every IVMI performance. But how can we measure the entity of this gap? And how can we intervene to reduce it? What we suggest is to embrace a larger perspective on performance practice, by means of applying scenographic theory to the domain of IVMIs.

In theatre, cinema and television, *scenography* relates to the study and the development of audio/visual, spatial and experiential composition of performance, by taking into account the perspective of two main stakeholders: performer and audience³. In the case of IVMIs, the dichotomy between performer's agency and audience connection is due to the constraints of immersive technology; nonetheless, analogous issues often arise when staging a play, filming a movie or broadcasting

³ Scenographic theory often includes the figure of the director among the stakeholders of a production; to make a simpler connection with IVMI performances we omit this detail, as in most cases the artist playing the IVMI is also the composer of the piece as well as the designer of the instrument.

a live show, even if completely different sets of technologies are in use. Shots of magicians or jugglers are quite challenging for example, as the presence of one or more cameras makes more difficult for these performers to disguise or show off their dexterous movements. A famous case is the one of the crystal-ball scenes in Jim Henson's 1986 movie *Labyrinth*, where artist Michael Moschen had to juggle blindly with only his right hand in the frame, while hiding the rest of his body behind David Bowie. In theatrical plays, it is the physical presence of an audience to challenge the performers instead. Actors are often compelled to face the spectators while interacting with stage props or with each other, sacrificing visual feedback/eye contact to create a sense of inclusion. One of the main aims of scenography is to account for these and similar scenarios. A well designed scenography fully immerses the spectators in the production, eliciting emotional and rational engagement [59] while seamlessly synthesising the performers' and the audience's experiences [44].

In the context of IVMIs, the scenography of a performance may be defined as the complete setup chosen to reproduce the instrument/VE on stage, make it playable to the musician and present it to the audience. This includes: (1) the technology dedicated to the immersion of the performer, like displays (e.g., HMDs, projections), tracking systems (e.g., head tracking, full-body motion capture, active/passive markers), physical user interfaces (e.g., joysticks, haptics) and sound monitors (e.g., headphones, speakers); (2) the technology addressing the audience's experience, like large screens, projected surfaces, lights and the power amplifier system; (3) the spatial arrangement of such technologies on stage, taking into account the freedom of action required by the musician to play the instrument, as well as size and position of the stage compared to the seat area or the parterre. In line with general scenographic theory, such a practice extends across design, curation and technical development [44].

When included in the design process of an IVMI performance, the development of a specific scenography may change how the VE is experienced, starting from disentangling immersive technology from the concept of *user*. Such a term is at the basis of most—if not all—conventional design approaches to VR, which tend to represent “the human subject as an omnipotent and isolated viewpoint” [30]. The great majority of IVMIs comply with this rationale. This is the main reason why, on stage, VR technology is almost exclusively employed to immerse the musician, creating the sense of isolation that we discussed earlier in this section. As opposed, musical VEs conceived for live performances would highly benefit from the exploration of novel design approaches, possibly discarding strict user-centric solutions.

In this context, scenographic theory may provide valuable guidelines to support experimentation. By dedicating time to the study and the development of a proper stage setup, artists may devise new ways of employing immersive technology, in compliance with the performer-audience paradigm that lies at the core of scenographic theory [59]. In other words, by definition a scenography has the power to turn the VR user into a performer, fostering an experience that is suitable for the entertainment of an external audience.

1.2 A New Approach to IVMI Performance

The design of a proper IVMI scenography is not an easy task though. In particular, the transition from user to performer proves to be a critical hit, as witnessed by the remarkable number of musical VEs that have been designed as instruments but never reached the stage. Indeed, despite the relatively large literature, very few works report the showcase of musical VEs in the context of concerts, for most IVMIs tend to be used as installations or research platforms, rather than as instruments for live performance [56, 35, 94, 47].

The aim of this chapter is to address this problem, and combine theory and practice to facilitate the design of IVMI scenographies. To do so, we propose a set of *dimensions* specifically conceived for the analysis of immersive stage setups. Such dimensions form a set of evaluation criteria that reflect the twofold nature of IVMIs. They stem from the detailed examination of the specifics of immersive VEs, combined with the practicalities of live music performances and in particular of those featuring DMIs. As a consequence, their application allows to extrapolate from a chosen stage setup the technical characteristics that affect these factors, and to qualitatively evaluate their individual impact on the showcasing of a generic IVMI performance. Furthermore, when the stage setup is coupled with a specific immersive instrument, the outcome of the analysis provides quick metrics to assess the experiential gap that likely divides performer and audience, also highlighting the main causes of such a disconnection.

It is worth noting that the scope of this work extends beyond the domain of VR. As detailed in the following sections, virtual performances and scenographies often span augmented and mixed realisations too, including see-through visual displays and a combination of physical and virtual stage props. This is the reason why we are referring to *immersive virtual* musical instruments as opposed to *virtual reality* musical instruments (commonly referred to as VRMI [78]), the latter being for the most part a sub-category of the former.

The actual relationship between these two classes of musical devices appears clear in the context of the categorisation of interactive environments proposed by Milgram and Kishino [65]. The two authors introduce a single continuous axis—the “virtuality continuum”—that goes from real environments (where everything is physical) to VEs (that host synthetic elements only), and encompasses in between all kinds of environments that mix physical and synthetic entities. On such a continuum, VRMIs belong to the far end of the spectrum (“virtuality”), and are distinct from devices that rely on technologies that lie closer to the “reality” end, like for example Augmented Reality (AR). However, the authors point out that virtual, augmented and any other kind of mixed technology can be characterised by different levels of immersion, regardless of their location on the continuum (see Section 3.1 for a more thorough discussion about immersion and its degrees of execution). In line with this perspective, IVMIs do not belong to a single point in the continuum, rather they cut across the spectrum; they include VRMIs, AR musical devices and any mixed solution in between, provided that the design of the instrument targets immersion.

Before moving forward, we'd like to remind the reader that this chapter is an extended and revised version of a pre-existing work, published by the authors in 2014 [12]. The decision to try to improve our contribution to this challenging research and artistic field comes from a very practical consideration. Over the seven years following the original publication, VR technologies have settled in the world of videogames and consumer electronics, and today the result of this process is the emergence of a new generation of immersive instruments and performances. For the first time, musicians have access to commercial IVMIs alongside more affordable and reliable resources for do-it-yourself development, to make new music and engage with new audio/visual experiences. And as expected, this is happening both inside and outside music studios. Companies are teaming up with underground as well as mainstream artists to popularise the usage of new immersive devices in performance settings, starting the exploration of innovative stage technologies to sell on the market of music and entertainment. In this vibrant scenario, the need for guidance in performance and instrument design is stronger than ever. The way we try to fulfil this need is by presenting a new set of cross-domain dimensions; by doing so, we aim at combining in one single critical perspective the practical—as well as cultural—implications that derive from the latest development of immersive musical technologies.

In line with this purpose, the rest of the chapter is structured as follows. Sections 2 and 3 discuss the main technological as well as experiential factors that play a role in the context of DMI performance and of immersive VEs, respectively. We will refer to these factors as *constraints*, a term originating from human-computer interaction [68] yet widely used in both DMI and VR literature [39, 93, 20, 32]. In particular, we embrace Magnusson's take on the subject, which deems constraints complementary to the *affordances* of an artefact/system [55]; in the context of this work, this means that by following cultural conventions and by adhering to technical and psychophysical requirements, it is possible to express at best the potential of DMIs, VEs as well as IVMIs. Starting from such constraints, in Section 4 we provide a detailed presentation of the set of dimensions we conceived to support the practice of scenographers and IVMI designers. Then, the following two sections exemplify how the dimensions may be applied to real-case scenarios. Section 5 analyses an assorted selection of IVMI performances spanning the last 30 years, with the aim to assess the type of experience provided to musicians and audience across all dimensions; while Section 6 shifts the focus on the future of immersive scenography, as we introduce novel stage setups and we use the dimensions to frame their potential when combined with IVMIs. Some of the solutions discussed in these two sections provide concrete examples of how to bring a musical VE to a live stage not only by using immersive VR technologies, but also by combining AR equipment/paradigms within the setup. Finally, conclusions are drawn in Section 7.

2 Constraints of the Digital Live Music Experience

DMIs are flexible tools that allow for the exploration of original musical and design practices. The vast potential granted by digital technologies makes it possible for designers and players to embrace the most daring sensing and interaction techniques, and to combine them with sound synthesis technologies that can also extend into the analog domain [76, 61]. Moreover, any mapping between musician's gestures and sound parameters can be devised almost arbitrarily, removing further limitations from the creative process [45, 91].

Unfortunately, this design freedom leads to great challenges when transferring a DMI from the studio to a live stage⁴. The type of musical exploration afforded by DMIs often manifests itself through bizarre and do-it-yourself equipment, unusual gestures, abstract sounds and idiosyncratic mapping between them. If not properly contextualised (in both a broad and a literal sense), these very distinctive features may impinge on the audience's experience of the show, as well as on the technical and expressive proficiency of the performers.

This section looks at DMIs from the perspective of live music performance. In particular, we discuss what we consider the main constraints linked to this form of expression/entertainment. Although centered on novel digital technologies, the list includes constraints that may as well inform the design of performances for traditional instruments. However, their overall impact is far more relevant when relocated within the domain of DMIs.

2.1 Stage Performance

On stage, performers need to be comfortable with their instruments. In an ideal scenario, a DMI plays the same regardless of *where* it is played, allowing the musician to build a live performance around the same affordances explored in studio and rehearsal spaces. Unfortunately, this is not always the case. Some DMIs are big and complicated, composed of parts that are difficult to assemble/disassemble or simply fragile. When dealing with such designs, the way the instrument is set up on stage often differs from the original studio configuration, forcing the performer to adapt their playing postures or even to sacrifice important visual/sound/haptic cues. Other musical systems impose requirements on the specifications of the stage itself. These include peculiar lighting, accurate microphone placement and support for multichannel audio playback; sometimes calibration procedures are required too, as in the case of multichannel audio or motion capture. If any of these elements is missing or merely not compatible with the DMI, some of the musical affordances and functionalities the performer grew accustomed to may become unavailable, right

⁴ A great portion of the DMI literature discusses how detrimental the apparent lack of design limitations may be on the musical appropriation of the instrument too, but this is beyond the scope of our work.

before the start of the show. A notable example of this contingency is the opening concert of the NIME Conference 2011. In that occasion, Carles López had to rework his performance on the fly, since the adverse lighting conditions of the stage made a large portion of his Reactable unresponsive (Figure 1).



Fig. 1: López's playing at the NIME Conference 2011; despite the use of curtains and dimmed lights, the instrument's optical tracking system kept malfunctioning, forcing López to adapt the execution of his pieces to the adverse situation. Image courtesy of Alexander Refsum Jensenius.

Approaching the stage unprepared clearly is a hazard and not all performers have the flexibility displayed by López (his performance was a success!). To avoid this risk, it is not uncommon for DMI musicians to organise live events in their practice studios, leveraging the very spaces where the instruments were designed, built and tested [97]. Yet, the appeal of a real venue is invaluable. Creators of a musical performance involving DMIs should dedicate particular attention to the phase of stage setup. Issues and necessities should be anticipated with care, from the most general and basic ones to the most specific and complex ones: will cables be long enough to connect the required hardware across the stage? How long does it take to calibrate and prepare the instrument? Are any of the pieces of equipment employed in the design difficult to install/use on a regular stage? Similar questions should arise early in the DMI's creation process, and could very well affect its design and behaviour.

2.2 Communication Between Performers and Audience

Communication between performers and audience is another fundamental aspect of musical performances. Often coupled with *cognition*, communication is one of main terms in use in music psychology and emotion research to frame the experience of playing and attending a live show. In this context, communication refers to the musician sending encoded messages for the spectators to be interpreted; more specifically, these resolve into music as well as related actions conceived to trigger specific emotions in the listener/viewer [48, 71]. However, authors like Gurevich, Treviño and Fyans thoroughly discussed how the application of such a model in the domain of DMIs is quite controversial, as it does not account for experimental and improvisatory music (to name a few), nor for the non-instrumental/intellectual engagement such instruments seem to be better suited for [38, 37].

To escape diversions from the main topic of this chapter, in this work we adopt a definition of *communication* that is more akin to Bonger's discourse on human-machine interaction [17]: non-verbal and non necessarily musical cues that define the interplay existing between audience and performers. From this perspective, communication can be analysed as a constraint, rather than as a yet-to-understand factor of music cognition.

Nonetheless, the way such a constraint is dealt with when designing a performance can still deeply influence and characterise the live music experience. Musicians should be able to perceive reactions of the audience, in order to adjust their playing and get a feeling of the ambience. For example, an improvised section could last longer or be cut short based on the hints performers can get from the audience. Or in large venues, performers could feel like getting physically closer to the spectators, or move around the stage also based on non-verbal cues. Spectators can communicate actively their emotions and appreciation to the performers via social and cultural conventions too, for example through gestures, applauding, shouting. Symmetrically, spectators should perceive musicians' expressions, gestures and looks, which are part of their playing style and together with the sonic outcome contribute to outline the performance. To this end, stages for live performances played in front of huge crowds typically include big screens showing close-ups of the musicians.

Apart from direct interplay between the parties, Bonger describes also another type of communication, happening in the context of "performer-system-audience" interaction. Performer and audience can indeed use the very technology set up on stage/in the venue (the "system") as a communication channel beyond sound and music. In his work, he discusses performances in which multimodality and—in particular—VR technologies are leveraged by the musicians to provide visual and haptic stimuli to single spectators, as well as multiple members of the audience. Finally, this type of communication includes the case of participatory performances, with spectators being able to use the system to input content in the performance and share information with the musicians. Some examples are the use of text messages as both sonification and literal communication means [29], or votes on the preferred type of music [92].

2.3 Music Ensemble

In performances involving multiple musicians, group dynamics are an essential aspect of both the performers' and spectators' experience. In fact, the interaction between musicians among the ensemble when playing DMIs may differ from what happens with traditional instruments. Moreover, the audience difficulty in understanding musician's gestures might increase with an ensemble, as it may also be difficult for the spectators to understand who is doing what [64]. Collaboration modes in digital ensembles can be separated in cooperation, communication and organisation modes [9]. Cooperation modes, when concurrent or complementary, allow musicians to share parts of sound generation processes or even allow other musicians to play their instrument. These choices and changes can be highlighted for the benefit of the audience - or even of other musicians, if they are not involved in the sharing process. Communication modes, such as exchanging messages or gestures indications can also be amplified for the audience, as done in [51], since they can be less visible than with acoustic ensembles. Finally, organisation modes, which allow musicians to define roles such as conductor and groups within the orchestra, are usually obvious from spatial arrangement of musicians in acoustic ensembles, and might need to be reinforced for the audience of digital orchestras.

3 Virtual Environments and the Constraints of Immersive Experiences

Compared to the case of live digital music, the compilation of a list of constraints capable of informing how we experience virtuality may seem overwhelming. The design of most, if not all, live DMI performances targets the delivery of one or more musical pieces; and while the details of the chosen technological setups vary from performance to performance and from artists to artist, their employment on and off stage is always dedicated to supporting the recreation and the diffusion of the featured music (as discussed in the previous section). Conversely, the variety of applications and scopes of systems capitalising on VR is astonishing, spanning industrial design [46], psychological and physical therapeutics [43], military training [52] and—of course—musical applications, just to name a few. From this perspective, it is quite hard to pin point all the requirements of such systems and scenarios, and to address in a single discussion the contingencies relative to employed technologies and common practices.

Luckily, the literature in VR research highlights an overarching theme that is common to all VR applications, and that can be used as the lens through which to analyse the constraints affecting the users' experience of generic VEs. This theme is the search for *presence*. In particular, presence has been described as the psychological sense of “being in the VE” [84], a specific state of consciousness that ought to be experienced by VR users. In an optimal scenario, when a user feels “present”

in a virtual world, they act as if the environment were real, physically and emotionally engaged in the application. Therefore, presence is targeted by all designers of VEs, regardless of the specific scope of the application or the technical details of the system. Furthermore, the concept of presence is tightly connected to disciplines like physiology, perception and psychology [82, 60, 80], making carefully designed narratives, settings and tasks necessary for it to be triggered.

In line with this scenario, we can consider constraints of VEs all the technological and experiential factors that play a role towards the establishment and the preservation of the feeling of presence in VR applications. In this section, we gather and discuss such constraints, placing emphasis on the aspects that will have a particular significance when crossing the domain of musical performance.

3.1 Immersion

Immersion is a key constraint in VR. The term refers to the description of the technology used to make the user feel present in the VE [81]. Immersive VR applications are characterised by a combination of equipment and techniques, the most common being wide stereoscopic viewports, multimodal feedback, detailed graphics, high framerate and large tracking areas. Such an arsenal may sound quite heterogeneous, and in fact it is not trivial to design and combine all its components as functional elements of a robust global system. Yet, the effect that this class of immersive technologies has on presence is immediate and conspicuous, such that researchers used to identify their technical specifications as the main constraints of VR [20]. But, nowadays, other immersive features are deemed fundamental too. For example those pertaining the design of the content of the VE, and in particular of those details that grant a coherent perception of the virtual objects, the surrounding virtual world and the virtual representations of the body of the user. In technical terms, this translates into scale, perspective and alignment. On a cognitive level, this coherence relies on components such as place illusion and plausibility, i.e. the sensation that the place and events occurring are real [79]. In a musical performance context with 3D avatars, plausibility for example seem to be strongly linked to eye contact with the musicians [8]. Presence is also strengthened by virtual body ownership, i.e. when one perceives their virtual avatar body as their own.

As discussed in [25], the effects of all these technologies and techniques are highly interconnected with one another. Moreover, the absence or the misuse of any of them may produce immediate breaks in presence [19]. For example, in a poorly designed VR setup the user may end up pulling the cable of a tracking device, or may thrust their hand through a virtual object, revealing its inconsistency. Similar contingencies have both perceptual and physiological consequences on the users, which can be measured to determine the extent of the experienced loss in presence [86]. Hence, immersion often fulfils the role of a filter too. Equipment and design techniques can be employed to block unwanted stimuli that come from the real world surrounding the VR setup, and that are often collectively referred to simply as *noise*.

These include the touch of hard boundaries of the tracking space, like sensor stands and walls [16, 67], or even the voice of people conversing close by⁵. Properly filtering noise from a VR setup is not enough to avoid all breaks in presence, but it is a good practice to minimise those caused by external reasons [80].

The strong connection between immersion and equipment means that different VR setups are characterised by more or less pronounced immersive features, regardless of the actual applications run with them. Then, the overall feeling of presence experienced by the user will depend on the specific combination of the immersive setup and the immersive design features of the software. But the role of the equipment/setting is so prominent that often times VR setups are assigned labels hinting at their intrinsic level of immersion [75]. These labels range from *fully immersive*, like the case of consumer VR headsets available nowadays and composed of HMDs, head and hand trackers, to *non-immersive*, denoting monoscopic screens and general purpose input devices, like mice, buttons and joysticks⁶.

The case of *partially immersive* setups (also labelled *semi-immersive*) is particularly interesting for the purpose of this work. Most of these mid-tier solutions capitalise on stereoscopic monitors and stereo-projected screens, occasionally coupled with head tracking. The result is a window on the virtual world, whose size is proportional to the rendering/projection area. Hence, the smaller the window, the more likely for the elements of the VE to end up beyond the clear-cut boundaries of the visual display and disappear from sight—especially during manipulation or locomotion. This eventuality endangers presence and represents the main technological limitation of partially immersive setups. And similar risks affect AR applications too, which leverage setups belonging to the same class.

Yet, monitors and projected screens provide VR designers with the opportunity to seamlessly combine real and virtual elements in the virtual experience. For example, a large projected setup allows to perceive the real hands and body of the user literally inside the VE, along with the virtual objects that populate it (or virtual objects inside the real world, as in the case of AR). Moreover, real-world objects and props may be used to carry out virtual interaction, hence entering the domain of Hybrid Reality [65]. As a consequence, the overall level of immersion that is achievable when using partially immersive setups largely depends also on *reality-virtuality continuity*, i.e., the set of immersive design features aimed at generating a consistent perceptual connection between the real and the virtual world. We can consider reality-virtuality continuity as an extension of the triad scale/perspective/alignment, that entangles rendering and tracking with the physical properties of real space.

AR displays can be used to cover a range of setups from partially immersive, e.g., integrating a few virtual elements in the physical space, to almost fully immersive, e.g., placing users in a virtual room or on a virtual stage. Because the physical space remains visible in all these cases, AR inherits from usual performance conditions:

⁵ Interestingly, this happens both in the context of research experiments, during which the subject may hear comments from the investigators or other lab members, and in casual game sessions, when friends/observers support the immersed player.

⁶ Indeed, when you are playing a videogame on your laptop or console, you are actually experiencing non-immersive VR!

direct visibility of the performers and other spectators, visibility of one's own body. These setups may also amplify errors in scale/perspective/alignment provided by a 3D display. However, these displays may also restrict 3D interaction opportunities to a subset of the techniques described below if part of the physical environment remains visible, e.g., navigation in a VE might be perceptually confusing if it is not correctly designed.

3.2 3D Interaction

Immersive technologies and design features are not the only means to trigger a sense of presence. In most cases, being able to interact with the VE encourages the user to deem the virtual world they are immersed in as real, and to forget that the experience is actually taking place in a different physical space. In other words, interaction is another powerful ally in the search for presence. The argument supporting this approach is that the reality of experience is defined by functionality rather than appearances, hence the sense of “being there” in a VE is grounded on the ability to “do” there [34, 77]. This does not mean that in an interactive system immersive technologies are superfluous or even a waste of money/resources; rather, interaction can be considered a constraint working on a different level from immersion, and both can be combined to describe VR more in depth.

Interaction with the virtual world (what we also referenced as “virtual interaction”) consists of altering the state of 3D models that populate foreground and/or background of the scene. This paradigm offers quite different perspectives—and challenges—compared to the case of interaction with 2D widgets, text and icons; for this reason, the term *3D interaction* is often used to distinguish it from the “desktop metaphor” employed in traditional personal computer environments [42]. Existing research on 3D interaction discusses an assortment of techniques, usually classified using the following categories: *selection*, *manipulation*, *navigation* and *application control* (the latter pertaining to menus and other VE configuration widgets, as a 3D extension of 2D interaction). In this section, we focus on the first three categories, as they provide a greater variety of controls with higher dimensionalities and are better suited to frame musical interaction in VEs.

Selection techniques allow users to indicate an object or group of objects in the VE. They are essential as they precede all manipulation techniques (i.e., to indicate which object will be manipulated) and some navigation techniques (e.g., to select a point of interest that the user wants to inspect). Several classifications of selection techniques have been proposed. Among them, Bowman et al. [18] classify techniques according to the object indication method (occlusion, object touching, pointing, indirect selection), activation method (event, gesture, voice command) and feedback type (text, aural, visual, force/tactile). More recently, Argelaguet and Andujar [1] proposed a set of design variables which allows for describing selection techniques according to, for example, the selection tool and how it is controlled, the control-display ratio or the disambiguation mechanism to avoid multiple selections. Most

common techniques involve either a virtual ray/cone projected from the user through the environment or a virtual cursor/hand mapped to the user's hand movements.

Manipulation techniques allow users to modify the spatial transform of elements of a VE, namely rotation, scaling and translation. They can also be used to modify their material (albedo, texture and other shading properties) through virtual tools such as virtual paint brushes and 3D palettes. Other techniques focus on the modification of the shape of composite 3D structures or 3D meshes, in particular through virtual sculpting metaphors. A recent review of such manipulation techniques can be found in [63].

Navigation techniques allow users to move inside the VE. This translates in the discovery of new areas and details of the virtual world, often segueing into the selection and the manipulation of virtual objects. From a technical perspective, navigation consists of a real-time update of the user's visual feedback carried out by the rendering engine, which provides a consistent dynamic representation of all the 3D models that cross the viewport. One possible classification of navigation techniques was introduced by [54] and separates them into three categories. *General movement* comprises all exploratory displacements through the VE, for example flying or walking. The case of walking is of particular interest; this type of navigation supports natural locomotion, a solution that has a strong impact on presence [83] and whose effectiveness can be further enhanced by means of walk-in-place immersive technologies and design features [73, 67]. The second category is *targeted movement*, that includes all techniques for which the user defines a target position and orientation within the VE. These can be discrete, when jumping or teleporting, but also continuous with smooth transitions between positions, such as those proposed in the Navidget technique [40]. Finally, *specified trajectory movement* techniques allow users to define a path through the VE which is then followed with different degrees of automation.

In the context of musical expression and IVMIs, these categories of interaction techniques can and have been used for all types of gestures, including selection of components of the instrument, excitation/production of sound and modulation of sound parameters [21]. For example, in *Drile* [10] a virtual ray technique is utilised for selecting tools and nodes of musical trees, while in Maki-Patola's VR percussion instrument [56] virtual sticks are used to trigger sounds. Techniques from a same category can be employed for both discrete and continuous controls in IVMIs. For instance, 3D navigation in *Versum* [3] allows for continuously controlling the volume of sound sources placed in the virtual environment. By changing to a discrete navigation technique, such as teleportation, one could trigger presets of sound mix, eventually playing them in a rhythm.

From a scenographical point of view, 3D interaction techniques do not all offer the same level of transparency [66], meaning that the influence of the musician on the VE can be more or less difficult to appreciate [13]. Navigation in VEs may be easily perceivable through the movements of avatars or changes in viewpoint. Manipulation and selection techniques however, especially when they involve subtle gestures (e.g., button presses, joysticks, finger poses) or complex graphical tools (e.g., sculpting or selection disambiguation), can prove more difficult to perceive and understand.

This effect can be reinforced in the case of techniques which require spatial alignment between the physical musician, or their avatar, and 3D graphical tools, for instance in the case of virtual rays techniques. In fact, a correct perceptual alignment then requires a fully immersive setup, either in VR or AR, making partially immersive setups less suited for specific interaction techniques.

The 3D interaction aspect of IVMIs constitutes one of the dimensions that we propose and is described in Section 4.5.

3.3 Collaboration and Observation

VR is not always a solo experience. Collaborative and social VEs [7, 26] are subject of extensive study, that pertains the relation among two or more immersed users and has yielded a large number of questions and results. In this context, users interact or cooperate within a shared VE, for example to collectively design industrial products or to join a gathering from remote locations.

When users leverage direct *collaboration* to achieve a practical common task, a number of factors influences efficiency as well as the dynamics of personal interaction. Analogous to the feeling of presence described above, *co-presence* [24] can be defined as the sense of being together in the VE. It has been shown to depend to a large degree on avatar appearance, as more realistic avatars tend to elicit a stronger sense of co-presence, as well as on the level of cooperation required to complete the actual task [70]. Another important aspect of practical collaboration in VEs is *awareness* [4]. Awareness can be defined as the understanding of other users' actions within the virtual world, a concept that relates strongly to the issues of musical performances with digital instruments covered in the previous section. Once again, embodiment (i.e., the provision of users with appropriate body images) has proven to have a strong impact on awareness [5]. Yet, other visual cues have also been proposed, such as a representation of the view cone of each user, signalling what is in sight and where the individual focus is.

The VR literature also discusses the case of *observation* without direct interaction. Virtual public speaking has been studied to understand the user/speaker's emotional response when performing in front of virtual audiences (immersed observers), leading to applications in psychotherapy for social phobias [85]. Moreover, other experiments focused on the observers themselves, and on the levels of presence and arousal triggered by watching virtual interaction as carried out by other users, using both immersive and non-immersive setups [50, 22]. As expected, these studies suggest that the lack of active involvement makes observers feel less engaged and less "present" in the VE compared to users. However, when both users and observers are properly immersed, witnessing real-time 3D interaction showed the potential to trigger a powerful perceptual experience, along with emotional responses way beyond the standards of non-immersive media and applications.

4 Dimensions of IVMI Scenographies

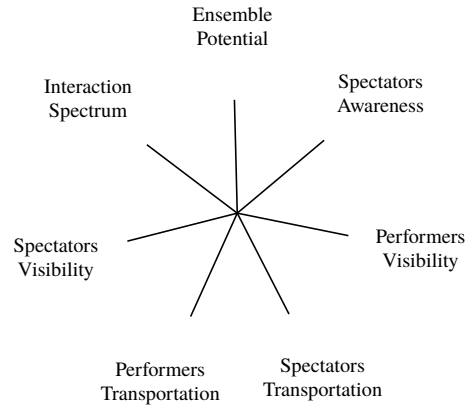


Fig. 2: Dimension space to describe performance setups of IVMIs

When designing and showcasing an immersive performance, artists have to take into account the full set of constraints that govern the experience of both digital live music and virtual interaction. Choosing the most appropriate stage technology to address each constraint may seem the obvious *modus operandi*, yet in a realistic scenario this straightforward approach reveals hard to apply. In first instance, some DMI and VR constraints appear to be orthogonal, meaning that good design in one domain tends to break constraints in the other. In other words, a technical solution specifically designed for musical purposes may end up hampering the device's VR functionalities and, vice versa, efforts targeting the virtual experience often degrade the pleasantness of the musical performance. Furthermore, constraints from different domains can combine, making standard technologies and common practices suddenly less effective in preserving engagement and expression.

For instance, moving to a VR audience-performer scenario, immersion deeply affects both the audience's and the musician's experience. An immersive performance acts on the audience's feeling of presence within the VE used on stage. As a consequence, the virtual instrument and all its 3D graphical components can be perceived, to a certain extent, as "real". In more practical terms, HMDs, single-user projections and head-tracking grant to the performer the level of immersion required to master the instrument, but exclude the spectators from the VE and cut direct communication between them and the performer. And the higher the performer's immersion (i.e., the more refined 3D musical interaction), the less intense the audience's experience (i.e., the less understanding and communication). As discussed in Section 5, the reverse is also true.

In this section we define the seven dimensions of performance setups of IVMIs and how they relate to the musical performance and VR constraints defined above. These can be visualised as a dimension space, as shown in Figure 2.

More than instruments based on physical, gestural or 2D graphical interfaces, IVMIs may create a strong asymmetry of performance experience between musician and spectators, depending on the display and interaction technologies used on each side. In turn this also generates different constraints, which we take into account by dedicating some dimensions to the audience experience and others to that of the performer's. For instance, the following dimensions focus on the performer's experience: *Performers Transportation*, *Ensemble Potential*, *Interaction Spectrum*, *Spectators Visibility*. Those targeting the audience experience are *Spectators Awareness*, *Spectators Transportation*, *Performers Visibility*. By placing them on the two sides of the dimension space shown in Figure 3 one can quickly judge the asymmetry in a given performance setup. The dimension space also distinguishes between interactive aspects, in the top half of the diagram, while immersion aspects are left to the bottom half.

The 7 dimensions emerged from multiple iterations and numerous discussions, with the aim of being usable for both the design and analysis of scenographies, addressing all aspects of the audience and musicians experience through the technical choices of performance setup.

4.1 Performers Transportation and Spectators Transportation

Performers Transportation and *Spectators Transportation* relate to the manner in which performers and spectators are immersed in the virtual musical environment, and to the extent to which the virtual and physical spaces intersect in a meaningful performative fashion. In particular, it indicates if the virtual stage is integrated in the physical space (or if it is surrounding it) and whether the setup is adequate to play/showcase the chosen IVMI.

It includes the following (non-exhaustive) range of technological settings to display the VE:

- a single monoscopic (2D) screen
- a volumetric display in the centre of a physical stage
- a mobile / handheld augmented-reality display
- a stereoscopic screen without and then with head-tracking
- a CAVE or set of stereoscopic screens
- an augmented-reality headset
- a virtual reality headset

Beyond visual displays, transportation also applies to auditory feedback (ranging from a monoscopic speaker, to ambisonics and binaural spatialisation) and to haptic feedback, including passive solutions (like the grips on the physical controllers required to play the IVMI) as well as proper actuators (ranging from a small vibrotac-

tile wearable, to exoskeletons for large-scale kineasthetic feedback). While targeting the enhancement of the feeling of presence within the VE may help, to achieve a high *Performers Transportation* these technologies have to be combined to allow the musician to play their IVMI on stage with no extra effort, compared to practice sessions carried out in a dedicated studio/lab space. Likewise, the ultimate scope of the *Spectators Transportation* dimension is to quantify to what extent the proposed *musical* experience feels real, and whether the display of musicianship is perceived as genuine as in the case of a traditional concert setting. It should be clear that the transportation dimensions are linked to, but do not overlap with, the constraint of immersion. They specifically highlight how the physical and virtual spaces intersect, similarly to what was proposed by Benford et al. in the context of shared VEs [6]. Although partially accounting for the need to feel present in VE, these dimensions incorporate all the stage performance requirements discussed in Section 2.1 and extend them in the domain of virtual worlds. The very term “transportation” was chosen to emphasise the focus on music, which is deemed capable on its own of psychologically transporting audiences into narratives, stories and fictional worlds [87, 28].

Transportation deeply affects both the audience’s and musician’s experience, yet in different manners. As a result, its measure tend to be highly asymmetrical. A straightforward example may be a scenography where the spectators are wearing VR headsets while the performer uses only a monoscopic screen—or the other way around as seen in *The Sound of One Hand*. Such a difference between the experience of the two stakeholders may not always be detrimental. For example, it is hard to imagine high *Performers Transportation* in the absence of interaction. Nonetheless, VEs that do not include interactive 3D objects, but are capable of physically reaching and surrounding the audience, sensibly enhance the transportation of the spectators [95]. More in general, HMDs, single-user projections, head-tracking and active/passive haptic feedback are all elements capable of granting the level of transportation required by the performer to master the instrument and play it on stage; yet, their use may exclude the audience from the VE and cut direct communication with the performer, unless the *Spectators Transportation* level is comparable. This translates into strong crossovers between transportation and other dimensions, such as *Performers Visibility*, *Spectators Visibility* and *Spectators Awareness*. For instance, VR headsets, which likely result in a high transportation value, impose a mediated view of musicians and spectators, e.g., with a 2D or 3D live-capture integrated in the VE, which in turn may reduce their visibility. On the other hand, with a low transportation level for the audience, their awareness might be constrained by the impossibility to visually align the virtual components of 3D interaction techniques, e.g., a virtual ray, with the physical hands of the musician.

4.2 *Spectators Awareness*

This dimension describes how well the audience perceives the virtual and physical interactions performed by musicians on the virtual instrument, i.e. the relation between their gestures, the instrument and the resulting changes in the sound.

It can be low for example when a technique such as virtual ray is used for the selection of distant parts of the instrument but this ray is either not visible at all for the audience or not visually co-located with the performer's hand. It can also be low if some physical interactions, e.g with physical sensors, are not visually reflected in the VE and the *Spectators Transportation* dimension is high.

The problem of abstract interaction is not unique to IVMIs, yet its occurrence is intensified by the employment of immersive technologies. Much like the case of IVMI performances, spectators are often incapable of fully grasping the workings of non-immersive DMIs, nor a causal relationship between action and music. As a result, the performance runs the risk to become opaque [33, 14], even confusing [27, 89], reducing the attributed agency [13] for the audience and in turn potentially degrading their experience [23]. DMI research suggests that this is due to the very metaphor of the instrument [33], as designs favouring intellectual and cognitive skills (e.g., live-coding environments, algorithmic devices) prove more prone to trigger abstraction compared to those leveraging familiar physical gestures [37].

When *Spectators Awareness* is low, the articulation between perceived manipulations, i.e. gestures and interaction techniques, and effects, i.e. controlled sound parameters, is not visible enough and there is risk of IVMIs being seen as secretive or magical instead of expressive [74].

In some cases, a breach into the virtual world is provided by means of screens that display the point of view of the performer. This solution may help the audience's understanding of the performance. Nonetheless, as explained in depth in Section 5, much is still left to imagination and interpretation, reason being that the IVMI is made visible but to the eyes of the spectators is not immersive (i.e., it does not surround the audience, nor the performer).

In cases where the transportation has a different value for spectators and performers or if the interaction techniques are too subtle or too complex, it is also possible to provide dedicated visual representations of the interactions for the audience. The design of these representations should however be chosen carefully. A correct balance needs to be targeted, between too little information, which results in a degraded subjective comprehension and potentially degraded experience [23], and too much information, which can lead to perceptual and cognitive overload. In the case of individual VR headsets or shared views of the VE for the audience, this level of detail can be interactively chosen by spectators [23].

4.3 *Performers Visibility and Spectators Visibility*

Performers Visibility and *Spectators Visibility* correspond respectively to the level of perception of the musician(s) by the audience and to the level of perception of the audience by the musician(s). It may take the following (non-exhaustive) values:

- not visible at all
- partially (from behind, from the side, with occluded parts)
- seeing fully in a simplified manner
- seeing a detailed 3D reconstruction or facing the physical performer

This dimension has a strong impact on the performer-spectator non-verbal communication.

Many commercial IVMIs and frameworks for immersive performances make use of avatars to represent the spectators. These are usually simple and can be chosen by users. They will therefore range between a medium and medium/high visibility levels depending on the level of detail provided on the appearance, behaviour and reactions. In these setups, performers often have more detailed or more expressive visual representations than spectators.

In a setup with lower *Performers Transportation* or *Spectators Transportation*, i.e. where the IVMI is integrated in the physical space, the physical spectators and performers can be seen more clearly if they are facing each other, with the instrument displayed between them.

4.4 *Ensemble Potential*

The *Ensemble Potential* dimension describes the ability for the scenography to accommodate multiple IVMIs or performers.

It is low when the setup only affords a single performer, for example because a head-tracked stereoscopic display is used or because the virtual environment was designed to host a single instrument or performer.

Depending on *Performers Transportation* and *Performers Visibility*, a high *Ensemble Potential* means either that the physical space can accommodate multiple performers collaborating on the same or with different instruments, or that the VE allows for displaying and/or navigating in 3D amongst multiple IVMIs.

Scenographies with a high *Ensemble Potential* should also ensure a correct co-presence [69], for example with high values of *Performers Visibility* and can provide access to a variable number of collaboration modes [9]. This dimension also strongly relates to the inter-actors and distribution in space dimensions used as part of the dimension space proposed by Birnbaum et al. [15]. Inter-actors describes the number of musicians while distribution in space specifies how the instrument extends in the physical space, ranging from a small device to a networked instrument. *Ensemble Potential* integrates both aspects, since IVMIs can virtually expand to integrate all musicians in a single shared VE.

4.5 Interaction Spectrum

This dimension describes what range of interaction techniques is permitted by an IVMI performance setup. These include the three categories of 3D interaction techniques described above, to which we add physical manipulations, i.e. musical interactions performed in the physical rather than virtual space.

3D selection techniques enable various types of musical gestures [21]. Although they would typically be associated with selection gestures, i.e. picking a component of an instrument, they can also serve as excitation, i.e. generating sound, or modulation gestures, i.e. changing the properties of the instrument. In fact, entering an object with a virtual ray may be used as an instantaneous excitation gestures to trigger a note, as done by Maki-Patola et al. [56]. It can also be used for continuous excitation when for example dragging a virtual cursor across the surface or inside the volume of virtual objects. In the context of public performances, selection techniques based on virtual rays require a high continuity (from physical hand to virtual ray) to be understandable by the audience, while virtual hands/cursors might be more tolerant (they are by definition not co-located when doing distant selection) and image-plane selection requires no (non co-located) visual feedback.

3D manipulation techniques can be used for both excitation and modulation musical gestures. Spatial transformations offer continuous controls from the changes in position, orientation and scale, but also discrete controls, which can be used i.e. as instantaneous excitation gestures, from collisions and intersections. Modification of appearance and shape can also serve as modulation gestures. For example, the *tunnels* of the *Drile* instrument [10] are 3D sliders which allow musicians to set the graphical parameters and associated sound parameters of 3D nodes of musical hierarchical structures. [66] proposes virtual sculpting as a way of setting musical parameters associated with the shape of a 3D mesh. Manipulation techniques can be distant, e.g. with 3D tools, or co-located, e.g. in the case of virtual sculpting. In both cases however, the musician's actions and the causal link between manipulation and musical result [13] are made visible to the audience by the visual changes in manipulated objects on which the focus is put. Therefore the lack of real-virtual continuity in manipulation techniques might not affect the spectator experience as much as other interaction categories.

3D navigation techniques can be used for most types of musical controls. Processes and parameters can be discretely selected before modification by entering associated volumes, such as the virtual rooms used in *Drile* [10]. Modulation of musical parameters can be achieved through displacement in parameter spaces, either continuous with general movement techniques or discrete with targeted movements. In the same manner, excitation gestures can be achieved by mapping relative position to virtual objects to the volume of associated sound processes, as done in *Versum* [3]. The impact of real-virtual continuity in the audience experience of 3D navigation depends very much on the granularity of the musician's position mapped to musical parameters. If the mapping is done according to the musician's movements within the space physically navigable, meaning that the user can physically walk to move through it, the audience understanding of the musician's impact on the sound will

require a high level of real-virtual continuity. However, if the navigation moves this physically anchored space in the VE, then the performed action is directly visible to all spectators from changes in the environment only and real-virtual continuity is not as necessary.

Physical interactions constitute another category of interaction which can be made available by a specific scenography, and corresponds to controls performed in the physical space, e.g on a control surface or an acoustic instrument. In order not to degrade the *Spectators Awareness*, these controls also need to be represented in the VE using changes in the performer's avatar or in the instrument appearance for example. Physical controllers and instruments can also be captured and rendered inside the VE.

5 Case study: Analysis of IVMI Performances

In this section we use the seven IVMI scenography dimensions to analyse different performances, and discuss their setups. This allows for practical observations on scenography, and their possible variations. The performances are introduced chronologically: the section tries to give a sense of evolution and change of the medium over time, both in terms of ideas, implementation, technology and diffusion. Performances have been selected giving precedence to pioneering solutions, and preferring well documented acts, both in the literature and on the web at large.

A visual representation of the analysis is given in Figure 3, in the form of a dimension space that provides a quick overview of each performance's properties. As mentioned in Section 4, it is structured both vertically and horizontally in order to provide a quick idea of the distribution of a scenography, between interaction and immersion and between spectators and performers.

5.1 Approaching a Performance Analysis

The analyses featured throughout this section start by dissecting the essential aspects of each performance. The main objective at the beginning of the process is to isolate the atomic components defining the stage setup, the IVMI, its use, and the expected behaviour of the performer(s) and audience. If the venue has some other peculiarities, it is also helpful to address them at this stage. As an example, the following are all valid questions which arise when starting the analysis process: are there HMDs involved? who is wearing them: the audience, the performer(s) or both? Is there a screen dedicated to the audience, how is it oriented? Is it hiding the performer from the audience, or vice-versa? Beyond the visual aspect, other important questions inquiry about the performance itself: how many performers are playing? How do the virtual and real instruments used on stage work? Are they easy to understand, or hindered by some design choice or technical limitation? Finally, our focus may shift

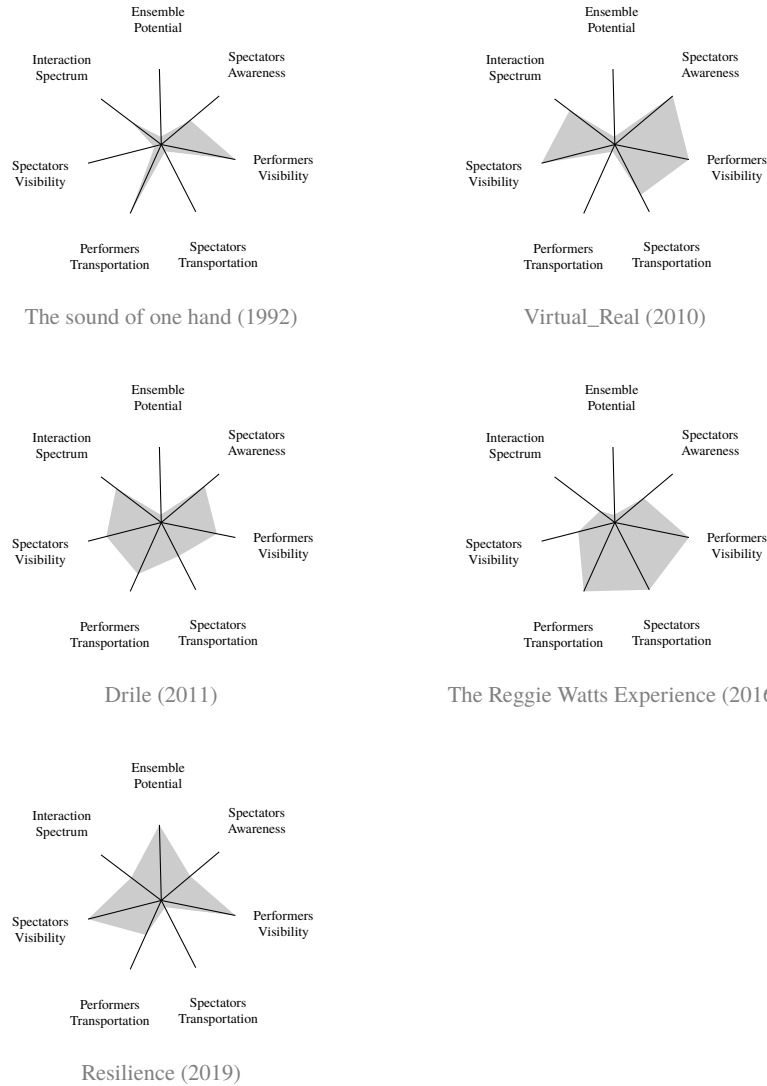


Fig. 3: Dimension Spaces for the analysed stage setups.

to the location: is everyone in the same physical location, or does the performance setup involve some form of telepresence? How good is the continuity between virtual and real elements on stage? What about the venue, seen from the performers' point of view? Therefore, the first part of the analysis consists of making a list of all the prominent bits which make the performance *that* performance: the resulting summary is not necessarily a technical survey. On the contrary, it can be interpreted

as a synopsis of the IVMI and its stage setup, from where the actual constraints will emerge. The outcome of this summary is a quick reference to consult when evaluating the seven dimensions.

Once the fundamental pieces of the performances (and their setups) have been identified and summarised, it is possible to start discussing how they fit within the dimension space as a whole. As a preference, analyses here presented start by addressing *transportation*. *Performers Transportation* and *Spectators Transportation* act as a solid ground for building the rest of the evaluation: as specified in Section 4.1, they can easily influence other dimensions. They encompass multiple feedback channels (visual, auditory, haptic), even though they tend to gravitate towards visual feedback, which has typically a heavy impact on presence. Available technologies are also affecting this bias towards visual immersion. Nonetheless, other channels should be considered carefully when investigating these dimensions. After evaluating transportation, it is reasonable to consider *awareness* and *visibility* dimensions, which also depend on multimodal feedback. Finally, the remaining dimensions can be addressed prioritising their prominence and importance within the performance.

Plotting the dimensions is a process involving a subjective judgement, especially when it comes to choosing the exact values used to generate the dimension space. Nonetheless, the seven dimensions are designed to highlight the asymmetries and relationships existing between the different aspects of a performance. Such constraints exist independently from the chosen numerical values: this is where a careful analysis potentially moves from being mostly subjective to being descriptive of a set of existing relationships. Certain technological setups are currently intrinsically incapable of providing, for example, high transportation both on and off stage, as-is. HMDs tend to hinder visibility, projected screens can break continuity between virtual and real elements, thus affecting transportation, and so on. So, the descriptive viewpoint provided by an accurate dimension space is of great interest despite the exact values used to create the plot. A reliable IVMI stage overview can be used not only to understand and analyse an already staged performance, but also to monitor and guide the design of a new one. The performance designer(s) could address early on some limitations, e.g., if *Spectators Visibility* and *Performers Transportation* are both considered important for a certain performance, a real time video or point cloud representing the audience could be used to improve *Spectators Visibility* when the performer is wearing a HMD, while maintaining high *Performers Transportation*.

After having carefully populated the dimension space, an additional final step is that of finding a one-sentence description of the analysed performance. In this Section these short descriptions can be found right at the beginning of each analysis. This final touch has at least two objectives: it implies a review of the analysis process, and it guides the future reader by highlighting the performance core values.

5.2 The Sound of One Hand

Pioneering performer's immersion for fine control. Jaron Lanier's *The Sound of One Hand* [49] was performed for the first time at SIGGRAPH in 1992.

Multiple virtual instruments are used during the performance, with Lanier playing them in turn. Sounds and notes were generated by hand movements, as they were transmitted to the instruments using a Data Glove. The Data Glove was also used by the musician to move and reach the instruments, which were sparse all around the VE. Instruments are described by the musician as autonomous and sometimes fighting back. Lanier talked in detail about these instruments, addressing how they are created and also how they take inspiration, visually and sonically, from real world instruments⁷.

A Head Mounted Display (HMD) was used by the artist in order to immerse himself in the VE, and therefore access the virtual instruments. This creates a setting where the musician can clearly be seen by the audience throughout the whole performance. On the other hand, it is impossible for the performer to see the audience. On stage, next to the performer, a screen was used to display a 2D projection of his point of view. This grants the spectators an access to the VE.

The primary dimension of this stage setup is *Performers Transportation*: the use of the HMD allowed for the musician to perceive a consistent world all around him, and to have access to fine 3D controls. However, the use of the HMD leads to the absence of *Spectators Visibility*. Conversely, *Performers Visibility* is quite high: Lanier played on stage, right in front of the audience, but he was also free to move and rotate, yet partially hiding his gestures. *Spectators Awareness* is limited, since the VE and the musician were perceived by the audience as two completely separated elements, the former projected onto a screen, the latter moving on the physical stage, with no continuity between the two. Furthermore, the screen projection was 2D and it displayed the musician's point of view, resulting into an extremely low level of *Spectators Transportation*.

The *Interaction Spectrum* includes 3D manipulation and navigation. Lanier opted for a point-flying navigation technique, a choice motivated by the artist's will to have an unconstrained and skilful way to explore the VE.

The scenographic level of this pioneering setup is understandably constrained, and it mainly focuses on the musician and his interaction with the VE. About the instruments, Lanier himself states that "*They emerged from a creative process I cannot fully explain*", and describes them as not immediately understandable, and also difficult to play. However, showing spectators a 3D projection aligned with the physical position of the musician on stage would remarkably enhance the audience's experience, providing immersion and increasing gestures continuity.

An interesting note: according to Lanier's impressions, the asymmetry between *Performers* and *Audience Visibility* resulted in him feeling vulnerable on stage - as opposed to what might happen when using rare and expensive technology for a

⁷ <http://www.jaronlanier.com/instruments.html>

performance. So, for the musician, added this combination of dimensions generated a *"more authentic setting for music"*.

5.3 Virtual_Real

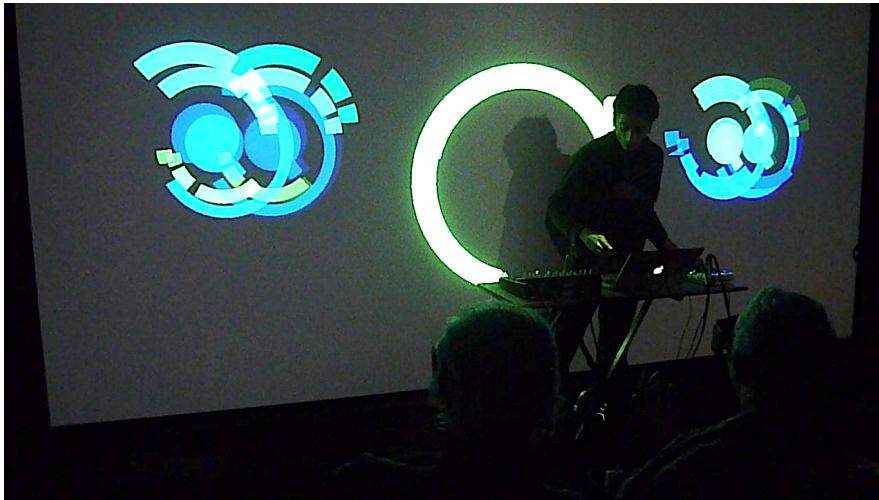


Fig. 4: USELESS_IDEA performing *Virtual_Real*, 2010. The shot frames two spectators wearing stereoscopic goggles, required to fully appreciate the hybrid virtual/physical stage.

Intense audience experience. The *Virtual_Real* performance [95] was born from a collaboration between Victor Zappi and the electronic composer USELESS_IDEA. It took place three times in Genoa, in 2010.

The performance was setup inside a laboratory room, which acted as an intimate venue. At the centre of the stage, the musician could play standard hardware controllers available in front of him. A single screen was positioned on stage, at his back. The screen displayed to the audience stereoscopic images of VEs populated by 3D objects, acting as both instruments and visuals. Thanks to optical motion capture the performer could move, touch or morph these virtual objects, in order to control audio effects. Thus, the setup allowed the musician to play both standard hardware controllers and non-immersive virtual instruments in front of the audience. 3D visuals and control algorithms were designed, tested and modified based on the artist's input and ideas. USELESS_IDEA played five tracks specifically composed for the event, each associated with a different 3D choreography.

Hardware controllers used by the performer included a laptop, a MIDI controller and a small mixer. The artist's dominant hand was tracked using passive reflective markers, allowing him to trigger interactions with the VE. The immersive content was designed to be experienced by the audience: despite the impossibility to provide headtracking for each spectator, the proportions between the projected screen size and the room size allowed the small audiences of nine spectators to enjoy a shared viewpoint, with no significant visual distortions. The audience could thus experience a stage where the performer, real items and virtual elements shared the same space (Figure 4).

This performance is strongly focused towards providing an intense audience experience. As a consequence, *transportation* is highly asymmetrical, with immersion affecting the spectators exclusively. The VE and its virtual instruments are perceived by the audience as coherently superimposed with the physical stage. This leads to a high *Performers Visibility*. *Performers Transportation* is absent, since the musician faces the audience and not the screen, while *Spectators Visibility* is high. *Spectators Awareness* is positively influenced by the possibility of clearly seeing the performer interacting with both real and virtual instruments. The musician's physical interfaces provide the same interaction transparency which could be expected in a traditional electronic music performance. Virtual instruments were coherently rendered with the audience's point of view, and the performer could be seen manipulating them. The sonic and visual result of such interaction were designed to be easily perceived. The *Interaction Spectrum* mainly included 3D manipulation techniques, with the performer moving and dragging objects around the VE scenes.

This single screen setup can create a strong involvement in the audience: virtual choreographies can be really convincing, and non-verbal communication with the performer can be really close to what would happen on a traditional stage. However, such an extremely audience-centric setup makes it impossible for the musician to use complex and potentially more expressive 3D interaction paradigms, thus limiting the possibilities of the virtual instrument. Slight setup modifications could generate a dual experience, in which the screen projection is dedicated to the performer, completely changing the scenographic outcome. The audience would no longer enjoy the perfect virtual/real environments consistency, while the musician would be immersed in the instrument, allowing for fine audio control.

5.4 Drile

Immersion for both ends. This performance was executed by Florent Berthaut in Bordeaux, 2011. The Drile instrument [10] used throughout the performance allows a musician wearing a stereoscopic goggles to execute live-looping in a 3D immersive environment. The performer uses handheld devices with pressure sensors in order to reach, excite and modulate the musical objects populating the environment. These objects are associated with the nodes of hierarchical live-looping trees, and their

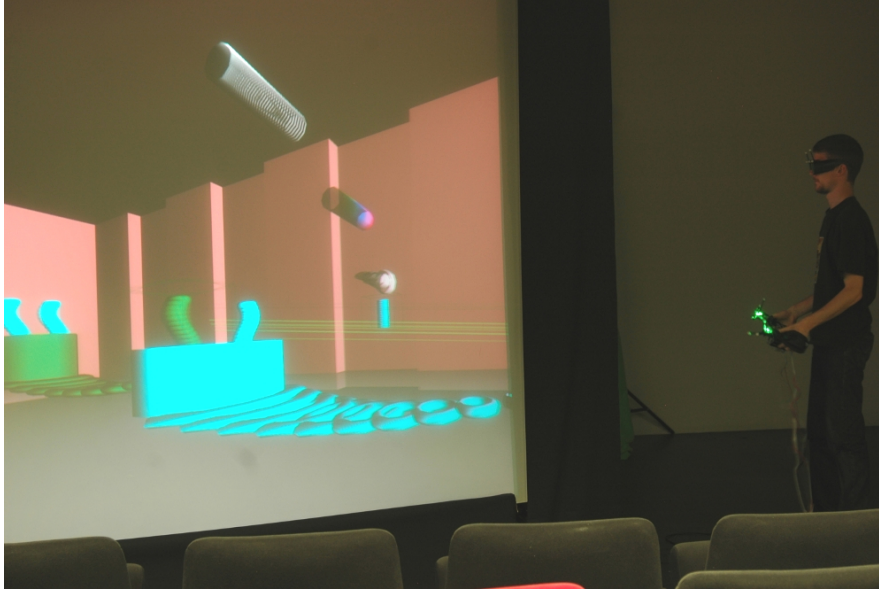


Fig. 5: Florent Berthaut performing with the IVMI *Drile*, 2011. The shot is taken from the seat area, and shows the 3D musical environment being pierced by the green virtual rays cast by the performer.

manipulation allows the musician to create and handle loops. Virtual rays are used to select and interact with the virtual objects.

Drile was shown on stage thanks to stereoscopic projections. Two screens, juxtaposed, were positioned on stage, with an angle between them. One screen was exclusively facing the performer, sideways. The second screen was rotated, so that it could face the audience. This arrangement had the screens defining an enclosed volume on stage. Therefore, both audience and performer perceived *Drile* as an instrument "contained" inside this volume (Figure 5). A correct perspective was granted to the performer by means of head-tracking, while a shared viewpoint was used to display the stereoscopic content on the audience screen.

This performance gives a highly symmetrical experience to audience and performer. *Transportation* is medium, since both parts can properly perceive the virtual instrument and the real stage while the virtual space is literally contained within the physical space. *Spectators* and *Performers Visibility* dimensions are quite high, since musician and spectators could directly see each other. *Spectators Awareness* is good, but hindered by the distance between the performer and the screen: virtual rays shown within the VE indicated which virtual objects the musician was manipulating, yet the instrument was operated standing one or more meters away from the screen. This distance breaks the continuity between the performer's hands and the

virtual rays. The *Interaction Spectrum* relies on virtual rays for the selection and manipulation of the 3D musical elements, but without physical manipulations.

This performance setup provides proper immersion for the audience and the performer, resulting in a great scenographic outcome and potential. Having a correct perspective for both parts allows the musician to have fine control of the instrument, and the audience to have a meaningful understanding of his actions. An alternative version of this setup could use a bigger, transparent screen dedicated to the audience. This screen would be placed between the spectators and the musician, allowing to overcome the absence of continuity between the performer's hands and the virtual rays shown in the VE. Since the musician's head and hands are tracked, additional visual effects and feedback solutions dedicated to the audience could be designed. This though could negatively impact *Performers Visibility*, and should be carefully implemented to avoid a negative outcome on *Spectators Awareness*.

5.5 The Reggie Watts Experience

A truly shared experience. This setup is based on the possibilities offered by social VR platforms. Users wearing a headset can share a virtual space, and interact with each other through 3D avatars. The performer Reggie Watts has been a recurring host of shows taking place specifically within the AltspaceVR platform, since 2016. His shows have been labelled as *The Reggie Watts Experience* and the performer keeps exploring the possibilities given by the format to this date .

Both the audience and the performer wear a HMD, which allows them to share the same virtual space and see each other. Reggie Watts movements on stage can be tracked thanks to full-body motion capture. He can use a microphone, controllers and effects, close to what he might do on a real stage. This kind of setup allows him to dance in front of the audience, see and address participants and move around the entire venue. The appearance of the avatars, venue and visual effects used throughout performances is designed to match the overall stylised aesthetic of AltspaceVR. Regarding the venues, different virtual spaces have been created and used, thanks to the possibilities given by the platform. Sometimes, visual effects can be seen, such as virtual fireworks, and simple, moving shapes. Tracking is available for the audience as well, based on the setup they have access to.

While wearing his HMD and tracking system, the performer can still interact with his own instrumentation, which is sometimes represented on the virtual stage by simple 3D models. AltspaceVR provides a tool which allows to host multiple instances of the same venue, so that countless number of spectators can participate at the same time. Each instance can host tens of participants, meaning that each member of the audience can be close to the stage. Spectators and the performer only see a limited part of the total number of participants currently present at the virtual venue. The completely virtual environment allows Reggie's voice and instruments to be spatialized, so that as he moves throughout the venue, it is clear for the audience where to look for him.

These performances focus on the idea of a shared space, and *Transportation* is strongly symmetrical: both the audience and the performer are immersed in the VE, as if they were physically present at the same venue. *Performers Visibility* is high, even when a huge audience is participating, thanks to the possibility of having multiple instances of the same performance, each hosting a limited number of spectators. *Spectators Visibility* is high, but only for those spectators which are in the same venue instance of the performer. So, from the performer's point of view, spectators are either really close and visible or not present at all. *Spectators Awareness* is limited to what can be understood from Reggie's limbs and body movements. Thus, audience experience mainly relies on his voice, music, posture and dancing. The presence of 3D models of his gear mitigates the limited awareness, for those cases in which the performer interacts with physical instruments. His avatar can in fact be seen bending over the controllers, making it easier to understand his posture in those particular moments. Regarding the *Interaction Spectrum*, virtual instruments are absent. 3D navigation is possible for the performer, and affects the spatialization of sound, but it's not used to interact with virtual instruments.

This performance allows a direct communication between the audience and the performer. Reggie can address his spectators, and interact with them. The possibility of seeing the performer moving, dancing and posing in the VE could be further explored, though. No virtual instruments are present, so the potential of the setup used to stage this performance is not completely explored yet. Virtual instruments could be added, which might be a way to create an even more compelling experience. Both the audience and Reggie share the same environment, and no perspective issues are present: this can overcome part of the limitations seen in more asymmetrical performances, and would allow a less constrained interaction design for virtual instruments. Nonetheless, the immediacy of having only the performer on stage has its own advantages: going to the extreme opposite could be detrimental to *Spectators Awareness*, and also negatively affect the spontaneous feel of the performance.

The Reggie Watts Experience is part of a set of immersive performances and Virtual Instruments which are exploiting the growing diffusion of consumer Virtual Reality setups. A variety of platforms is being developed, each addressing different scenarios: immersive music making, remote participation to live events, VR dance clubs and so on. *Electronauts* is a VR instrument for beat making and jamming. The creators also showcased an augmented/mixed reality video of a session where a performer playing the Electronauts instrument jams along with other musicians (guitar, sax drums). AltspaceVR is providing a platform for performers like Reggie Watts to create shared musical experiences, and other companies aim to provide similar setups. MelodyVR allows to capture and share immersive videos from live concerts, which can be experienced on a VR HMD. Online multiplayer videogames such as Fortnite have been used to host musical performances. Even if not immersive for audience or musicians, such endeavours show a growing interest in the exploration of novel possibilities in the field of virtual musical performances.

5.6 Resilience

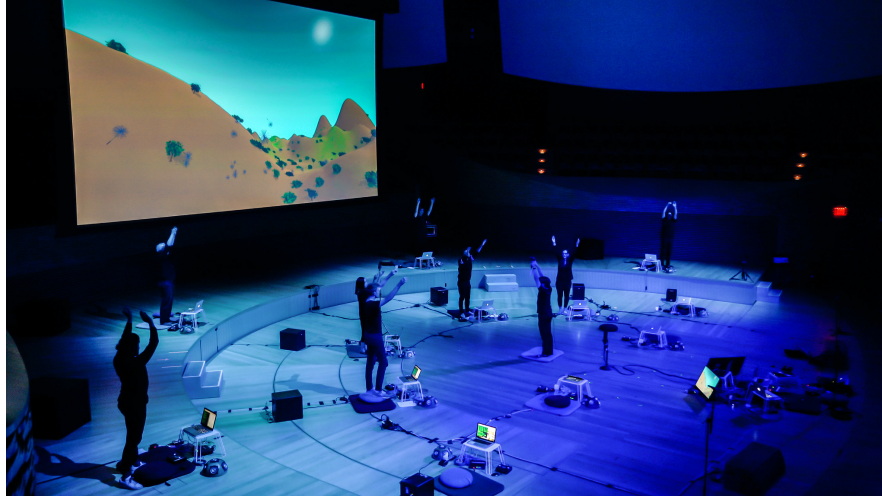


Fig. 6: *Resilience* immersive musical performance, 2019. The conductor stands at the center of the stage, wearing an HMD and leading the orchestra via both hand gestures and 3D interaction. Image courtesy of Ge Wang, Stanford Laptop Orchestra.

A laptop orchestra with a VR Conductor. This performance is designed for a laptop orchestra and one VR performer/conductor. *Resilience* [2] was performed in June 2019, at the *2019: A SLOrk Odyssey* concert at the Bing Concert Hall of Stanford University.

The VR performer is at the centre of the stage, wearing an HMD and acting as a conductor. Surrounding him, eight performers are positioned on two separate rings. The performer's hand movements are tracked by handheld motion controllers, while the rest of the ensemble has access to tether controllers. Each performer has a laptop and speaker array. The VR performer is facing away from the audience, in the direction of an oversized projection screen. Thus, the audience has a view of the conductor, the orchestra and a 2D projection of the environment experienced by the VR performer (Figure 6).

The performance was structured in three movements, with the VR conductor cueing the orchestra throughout the piece with his body movements. By using motion controllers, he sometimes also triggered flashes of lightning. The orchestra elements used tether controllers to affect the movements of virtual seedlings and their visual aspect, and excite synthesised sounds. The way the performers movements were acted out, and the timbre of the synthesised sounds changed with each piece movement. The whole ensemble at times also acted as a whole meta-instrument, per-

forming wave gestures which were paired with movements of a wind timbre across the ensemble. When this happened, the virtual seedlings changed their direction accordingly. During the entire performance, the point of the projection shown to the audience was curated by the head movements of the conductor, which the creators have thoroughly evaluated and rehearsed. The same 2D projection was rendered on small monitors available to the orchestra performers.

In terms of fruition, this performance provides different experiences to the audience, conductor and orchestra. *Performers Transportation* is high for the conductor, who is immersed in the VR environment, while the orchestra only experiences the VE on a small monitor. The conductor entirely misses the real stage, which conversely is the main space experienced by the other performers. Because of this, overall *Performers Transportation* can be considered of medium level. *Spectators Transportation* is limited: the stage is clearly in front of the spectators, while the virtual environment is displayed on a screen from the conductor's point of view. *Performers Visibility* is high for the audience and the orchestra, while the conductor can only perceive the virtual environment. *Spectators Visibility* is high for the orchestra, and absent for the performer. *Spectators Awareness* is positively influenced by the clearly visible choreographed movements of conductor and orchestra, which are affecting the sonic outcome of the piece and the visuals of the virtual environment, thanks to the tethered controllers. This performance *Ensemble Potential* is good, as the piece is designed for a conductor and orchestra. Nonetheless, co-presence is limited for the conductor, who cannot perceive their own orchestra if not sonically.

Resilience could be described as a carefully planned laptop orchestra piece with live visuals, featuring the addition of a conductor immersed in a VR environment. Audience access to the VE is provided through a 2D projection, curated by the conductor in real time. This can be used as an expressive channel, to the expense of audience immersion, which could otherwise be improved by introducing a stereo projection with shared point of view (see Virtual_Real and Drile performances).

6 Towards the Design of Novel Scenographies

The rapid growth of consumer and professional VR technologies is offering new interesting perspectives to IVMI designers and performers. As hinted by the analyses presented in the previous section, a fair amount of stage setups has been explored over the last 30 years, leading to extremely different experiences and related dimensional spaces. Nonetheless, there is still much to experiment with and to discover. Every year, technologies that once were seen only in research laboratories or during specialised scientific events become available in public and entertainment spaces, and some even populate the shelves of electronics stores. Some examples are the large immersive multi-projection systems now found in several museums and performance spaces, as well as the first wave of see-through headsets that hit the market just a few years ago. While facilitating the design of more advanced and more daring virtual experiences, these technologies embrace specifications that make them more

and more compatible with digital media and—in particular—audio standards⁸. As a result, the creative horizons of VR musicians keep expanding, thrust by the embedding of devices, materials and arrangements that had never before been available to convey musical expression in live settings.

In this section, we take the liberty to suggest three solutions that propose unique takes on the virtual musical experience and that, to our knowledge, are yet to be explored. It is important to remark that we are not going to describe scenographies *per se*, though. The technological and spatial composition of a virtual musical performance depends necessarily on both the instrument and the stage (Section 1.1), and refers to a precise instance (or a series of instances) of the show. As opposed, now we are about to discuss the use of immersive technologies in precise stage arrangements that encompass performers and spectators, yet without focusing on any specific IVMI or performance. In this context, the dimensions allow us to carry out an analysis of the *potential* of these stage setups, in terms of their ability to accommodate various categories of musical instruments and to create impactful scenographies for/with them. At this point, it should also be clear to the reader that no solution is perfect and the setups we are about to introduce are no exception.

6.1 Co-located Antithetical Immersion

We start with something relatively easy to achieve, at least from a purely technological perspective. Let us consider what appear to be two antithetical immersive solutions, in particular those used in *The Sound of One Hand* and *Virtual_Real* (Sections 5.2 and 5.3). The former features HMD and dataglove to give the performer full access to the VE (high *Performers Transportation* and wide *Interaction Spectrum*), though limiting *Spectators Transportation* and *Spectators Awareness*; the latter leverages exo-centric 3D projections, that convincingly merge virtual and physical world to the eyes of the audience (remarkable *Spectators Transportation* and *Spectators Awareness*), at the cost of *Performers Transportation* and *Interaction Spectrum*. Although often used separately (e.g., [36] and [90]), these immersive setups can be combined to balance out most of their individual shortcomings.

In practical terms, what we envision is a stage where headsets are used by performers⁹ and stereoscopic projections are designed for the audience. This scenario co-locates on the same physical stage immersive technologies that differ in structure and target, allowing to display the VE and interaction from two distinct perspectives simultaneously: the one of the performer (as rendered on the HMD) and the one of the audience (as rendered on the screen). This permits to reach high values of transportation for both performer and audience, and strong *Spectators Awareness*. Furthermore, such a setup provides access to all 3D interaction techniques, making

⁸ To corroborate this argument, we would like to point the reader to the beautiful *Alt F* performance showcased at NIME 2021, during which DMIs well-known to old time attendees of the conference were played for the first time in a fully immersive online environment!

⁹ As discussed in Section 3.1, headsets combine an HMD with tracking and input devices.

it compliant with a variety of IVMIs and leading to the design of scenographies characterised by a broad *Interaction Spectrum*. Unfortunately, the use of headsets makes the visibility dimensions quite asymmetrical, yet scales quite well in case of ensemble performances.

6.2 Augmented Worskpace and Spatial Paradox

The second setup that we propose promotes a rather “unorthodox” experience of the space that performer and spectators share. Right at the beginning of this chapter (Section 1), we mentioned the possibility to play with the scale of the VE in paradoxical ways, the most common example being virtual instrumentation that exceeds the physical size of the stage (e.g., [49]). Now we take a step in the opposite direction. We describe a solution to make music with a virtual world in miniature, that can fit the hands of the performer, but is still capable of surrounding a full audience!

The inspiration for this concept comes from artist Hicham Berrada and his work *Présage*. Berrada filmed a 360 view of the inside of a small water tank, while pouring coloured chemicals into it; he then scaled up the video to fit a large multi-projection installation, where spectators could experience a stroll at the bottom of the lively tank. Our take on this setup replaces the water tank with a medium to small sized VE, populated with musical objects and embedded in the performer’s workspace via an AR headset (like the Microsoft HoloLens). In a separate room, the audience is hosted inside an immersive stereo-projection system; here, the same VE is scaled up of two or more orders of magnitude and rendered as if the seat area/parterre were inside of it, facing the performer. Furthermore, stereo-cameras can be easily installed on both sides of the setup, so that the AR workspace could include a miniature volumetric render of the audience and the stereo-projections could showcase the titanic body and gestures of the performer¹⁰. The result is a paradox, a non-existing shared space where musician, spectators and virtual objects can be huge or tiny, depending on the beholder.

This unusual setup may support the design of scenographies that excel in most dimensions, in particular those pertaining transportation and visibility. The main drawbacks may though come in terms of low *Ensemble Potential* and *Interaction Spectrum*. Indeed, sharing the AR workspace between more than two performers may reveal problematic, while the overall spatial design aligns well only with specific interaction techniques and IVMI metaphors.

¹⁰ The seat area may be even virtually embedded in one of the movable parts of the instrument, turning the execution of a piece into a lively ride for the audience; this type of extremely embodied musical experience would likely cause a high occurrence of motion sickness, but who are we to define the boundaries of artistic expression?

6.3 Double Sided Virtual World

We conclude our review of proposed stage setups with a technically challenging yet visually impressive solution. The aim of this last entry is to employ a single VR/AR technology to immerse both performer and spectators, while they are physically present in the same venue. By leveraging the *Pepper's ghost effect*, an acrylic semi-transparent screen can be set up to obtain a double sided reflective surface, that splits the stage from the seat area and forms two distinct windows on the virtual world—one for the musician and one for the audience. The screen has to be installed at the edge of the stage with a 45 degree horizontal tilt, so that one of its sides leans towards the spectators. Then, two projection surfaces are placed above and below it; projections reaching the top surface are reflected on the audience's side of the semi-transparent screen, while projections directed to the bottom surface are reflected on the musician's side. Such a setup minimises interference between the two reflections, so that both performer and audience can use the screen to have a clean stereo-view of the VE, from their own perspective. The way the VE is rendered on the two sides may even differ in level of details or content! Furthermore, portions of the screen not reflecting any light maintain their see-through nature. This allows to include physical props within the VE or, vice versa, to augment traditional musical gears with virtual widgets.

In our 2014 work [12], we described a prototype scenography based on this double sided setup, built and tested in a VR laboratory. Despite the obvious advantages of working in a controlled setting as opposed to an actual stage, that experience highlighted the effort required to install the screen apparatus and to calibrate it along with a tracking system. Nonetheless, once in place the setup revealed quite remarkable capabilities. Both sides of the stage can support 3D visuals, tracking and multimodal feedback without interfering with each other, hence leading to very high peaks of *Performers* and *Spectators Transportation*. As previously mentioned, interaction is potentially extremely varied (wide *Interaction Spectrum*) and easy to understand (high *Spectators Awareness*), with the caveat that the playing of the IVMI must happen in the space between the musician and the audience. But where this setup excels are the visibility dimensions; thanks to the semi-transparent screen the physical bodies of both musician and audience can be completely visible to one another, much like the case of a traditional musical performance. The only clear limitation concerns *Ensemble Potential*, for the employment of such a complex projection-based setup makes it extremely difficult to immerse more than one musician on stage at a time.

7 Conclusion

In this paper, we investigated the scenography of immersive virtual musical instruments. We first reviewed the constraints of both immersive virtual environments and digital musical performances. From these, we derived 7 dimensions for the design

of scenographies of IVMIs. We finally demonstrated how this dimension space can be used to analyse past performances and how it can inform the design of new ones.

We also believe that this dimension space may result into an opportunity to improve the quality of IVMI scenographies. Scenographers may employ the dimensions to intervene on the most critical details of the stage setup, and choose technologies and spatial arrangements that make the performance as inclusive as possible, without the need to modify the instruments' metaphor. Moreover, the proposed approach to IVMI performance practice has the potential to influence instrument design too. For example, when the topology of a venue imposes too many constraints to build a proper scenography, the instrument designer may use the dimensions as a set of guidelines to adapt the IVMI to the encountered limitations.

In similar eventualities, the outcome of the dimensional analysis carries important design feedbacks, that might extend even beyond the specific stage scenario. Maybe the metaphor designed for the IVMI is simply too complex/idiosyncratic to result comprehensible to an external audience, whether or not the venue is suitable for the showcase of immersive performances; in other cases, it might be the specific combination of some of the parts of the instrument's metaphor to hinder the transition from user to performer—for example, an interaction technique that is not compatible with the chosen visualisation paradigm. So, another scope of the dimensions is to preventively foster this kind of analysis, and push the designer to question the nature of their musical VE (i.e., instrument or installation?) during the very design phase.

We can see multiple extensions of our dimension space, which would allow for: 1) a stronger integration of the various perceptual aspects in a performance and 2) refinements in the analysis to handle the complexity of performance scenographies. First, the dimensions that we proposed, in particular *transportation*, tend to focus on the visual immersion of both the audience and performers, i.e., the choice of display technology. While presence in a virtual environment and the experience of musical performances are very strongly impacted by the visual perception, other modalities are also essential. Our dimensions could therefore be refined to take into account the auditory and haptic transportation and the interactivity for the audience.

Second, in this paper, we chose to use the word scenography to describe technical design choices. In this regard, a possible refinement of our dimension space would be to distinguish between *stage setups*, which can be informed by the dimensions, and the development of the setups into shared musical experiences (i.e., actual IVMI scenographies!), which require further discussion, and potentially an even more qualitative analysis approach. However, given a set of choices, the diversity of potential implementations remains very high. In fact, the relationship between constraints and the outcome of a performance is even more complex and more counter-intuitive than what one would expect. Skilled scenographers may carefully pay attention to the direct consequences of design decisions across virtuality and music, yet the strong entanglement among the constraints (and the different stakeholders) may make any prediction quite inconsistent. For example, it is hard to suspect that replicating on stage the same exact setup used by the musician to rehearse in studio could be detrimental to the outcome of the performance. Such a design approach would preserve the intimacy with the instrument developed by the performer over hours of practice

(DMI constraint, Section 2.1) and it would reinforce the level of immersion that is achieved on stage (VE constraint, Section 3.1); yet, it may clash with how the actual IVMI lends itself to a live stage realisation, as well as with venue specifics, audiences' expectation and—always present—miscellaneous contingencies. As a consequence, the term “scenography” as intended in this work does not equal a predictable experience.

References

- [1] Argelaguet F, Andujar C (2013) A survey of 3d object selection techniques for virtual environments. *Computers & Graphics* 37(3):121–136
- [2] Atherton J, Wang G (2020) Curating perspectives: Incorporating virtual reality into laptop orchestra performance. In: *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp 154–159
- [3] Barri T (2009) *Versum: audiovisual composing in 3d*. Georgia Institute of Technology
- [4] Benford S, Bowers J, Fahlén LE, Greenhalgh C (1994) Managing mutual awareness in collaborative virtual environments. In: *Proceedings of VRST*, pp 223–236
- [5] Benford S, Bowers J, Fahlén LE, Greenhalgh C, Snowdon D (1995) User embodiment in collaborative virtual environments. In: *Proceedings of CHI 95*, New York, NY, USA, pp 242–249
- [6] Benford S, Greenhalgh C, Reynard G, Brown C, Koleva B (1998) Understanding and constructing shared spaces with mixed-reality boundaries. *ACM Transactions on computer-human interaction (TOCHI)* 5(3):185–223
- [7] Benford S, Greenhalgh C, Rodden T, Pycock J (2001) Collaborative virtual environments. *Commun ACM* 44:79–85
- [8] Bergström I, Azevedo S, Papiotis P, Saldanha N, Slater M (2017) The plausibility of a string quartet performance in virtual reality. *IEEE transactions on visualization and computer graphics* 23(4):1352–1359
- [9] Berthaut F, Dahl L (2015) Boeuf: a unified framework for modeling and designing digital orchestras. In: *International Symposium on Computer Music Multidisciplinary Research*, Springer, pp 153–166
- [10] Berthaut F, Desainte-Catherine M, Hachet M (2010) Drile: An Immersive Environment for hierarchical live-looping. In: *New Interface for Musical Expression*, Sydney, Australia, p page 192, URL <https://hal.archives-ouvertes.fr/hal-00530071>
- [11] Berthaut F, Desainte-Catherine M, Hachet M (2011) Interacting with 3D Reactive Widgets for Musical Performance. *Journal of New Music Research* 40(3):253–263
- [12] Berthaut F, Zappi V, Mazzanti D (2014) Scenography of immersive virtual musical instruments. In: *VR Workshop: Sonic Interaction in Virtual Environments (SIVE)*, 2014 IEEE, Minneapolis, United States,

- DOI 10.1109/SIVE.2014.7006285, URL <https://hal.archives-ouvertes.fr/hal-01222454>
- [13] Berthaut F, Coyle D, Moore J, Limerick H (2015) Liveness Through the Lens of Agency and Causality. In: Proceedings of NIME
 - [14] Bin A, McPherson A, Bryan Kinns N (2016) Skip the pre-concert demo: How technical familiarity and musical style affect audience response. In: NIME '16: Proceedings of the 2016 conference on New interfaces for musical expression
 - [15] Birnbaum D, Fiebrink R, Malloch J, Wanderley MM (2005) Towards a dimension space for musical devices. In: Proceedings of the 2005 conference on New interfaces for musical expression, National University of Singapore, pp 192–195
 - [16] Boletsis C (2017) The new era of virtual reality locomotion: A systematic literature review of techniques and a proposed typology. *Multimodal Technologies and Interaction* 1(24)
 - [17] Bongers B (1999) Exploring novel ways of interaction in musical performance. In: C&C '99: Proceedings of the 3rd conference on Creativity & cognition, ACM, New York, NY, USA, pp 76–81, DOI <http://doi.acm.org/10.1145/317561.317576>
 - [18] Bowman DA, Hodges LF (1997) An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In: SI3D '97: Proceedings of the 1997 symposium on Interactive 3D graphics, ACM, New York, NY, USA, pp 35–ff., DOI <http://doi.acm.org/10.1145/253284.253301>
 - [19] Brogni A, Slater M, Steed A (2003) More breaks less presence. In: Proceedings of Presence 2003: The 6th Annual International Workshop on Presence, URL <http://www.presence-research.org>
 - [20] Bryson S (1995) Approaches to the successful design and implementation of vr applications. *Virtual reality applications* pp 3–15
 - [21] Cadoz C (1999) *Musique, geste, technologie*, Éditions Parenthèses
 - [22] Calvert SL, Tan SL (1994) Impact of virtual reality on young adults' physiological arousal and aggressive thoughts: Interaction versus observation. *Journal of applied developmental psychology* 15(1):125–139
 - [23] Capra O, Berthaut F, Grisoni L (2020) All you need is lod: Levels of detail in visual augmentations for the audience. In: The 20th International Conference of New Interfaces for Musical Expression
 - [24] Casanueva J, Blake E (2000) The effects of group collaboration on presence in a collaborative virtual environment. In: *Virtual Environments 2000*, Springer, pp 85–94
 - [25] Cho D, Park J, Kim GJ, Hong S, Han S, Lee S (2003) The dichotomy of presence elements: The where and what. In: *VR*, IEEE Computer Society, pp 273–274, URL <http://dblp.uni-trier.de/db/conf/vr/vr2003.html#ChoPKHHL03>
 - [26] Chung D (2005) Something for nothing: understanding purchasing behaviors in social virtual environments. *Cyberpsychology & behavior* 8(6):538–554
 - [27] Collins N (2003) Generative music and laptop performance. *Contemporary Music Review* 22(4):67–79

- [28] Costabile KA, Terman AW (2013) Effects of film music on psychological transportation and narrative persuasion. *Basic and Applied Social Psychology* 35(3):316–324
- [29] Dahl L, Herrera J, Wilkerson C (2011) Tweetdreams : Making music with the audience and the world using real-time twitter data. In: *Proceedings of the International Conference on New Interfaces for Musical Expression*, Oslo, Norway, pp 272–275, DOI 10.5281/zenodo.1177991, URL http://www.nime.org/proceedings/2011/nime2011_272.pdf
- [30] Davies C (1998) Osmose: Notes on being in immersive virtual space. *Digital Creativity* 9(2):65–74, DOI 10.1080/14626269808567111, URL <https://doi.org/10.1080/14626269808567111>, <https://doi.org/10.1080/14626269808567111>
- [31] Davies C, Harrison J (1996) Osmose: towards broadening the aesthetics of virtual reality. *SIGGRAPH Comput Graph* 30:25–28, DOI <http://doi.acm.org/10.1145/240806.240808>, URL <http://doi.acm.org/10.1145/240806.240808>
- [32] Dickey MD (2005) Brave new (interactive) worlds: A review of the design affordances and constraints of two 3d virtual worlds as interactive learning environments. *Interactive learning environments* 13(1-2):121–137
- [33] Fels S, Gadd A, Mulder A (2002) Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound* 7(2):109–126
- [34] Flach JM, Holden JG (1998) The reality of experience: Gibson’s way. *Presence* 7(1):90–95
- [35] Gelineck S, Böttcher N, Martinussen L, Serafin S (2005) Virtual reality instruments capable of changing dimensions in real-time. *Proceedings Enactive*
- [36] Graham R, Bridges B, Manzione C, Brent W (2017) Exploring pitch and timbre through 3d spaces: embodied models in virtual reality as a basis for performance systems design. In: *NIME*, pp 157–162
- [37] Gurevich M, Cavan Fyans A (2011) Digital musical interactions: Performer–system relationships and their perception by spectators. *Organised Sound* 16:166–175, DOI 10.1017/S1355771811000112, URL http://journals.cambridge.org/article_S1355771811000112
- [38] Gurevich M, Treviño J (2007) Expression and its discontents: toward an ecology of musical creation. In: *Proceedings of the 7th international conference on New interfaces for musical expression*, pp 106–111
- [39] Gurevich M, Marquez-Borbon A, Stapleton P (2012) Playing with constraints: Stylistic variation with a simple electronic instrument. *Computer Music Journal* 36(1):23–41
- [40] Hachet M, Decle F, Knodel S, Guitton P (2008) Navidget for easy 3d camera positioning from 2d inputs. In: *3DUI '08: Proceedings of the 2008 IEEE Symposium on 3D User Interfaces*, IEEE Computer Society, Washington, DC, USA, pp 83–89
- [41] Hamilton R (2019) Collaborative and competitive futures for virtual reality music and sound. In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, IEEE, pp 1510–1512

- [42] Hand C (1997) A survey of 3d interaction techniques. In: Computer graphics forum, Wiley Online Library, vol 16, pp 269–281
- [43] Hodges LF, Anderson P, Burdea GC, Hoffmann H, Rothbaum BO (2001) Treating psychological and physical disorders with vr. *IEEE Computer Graphics and Applications* 21(6):25–33
- [44] Howard P (2019) *What is scenography?* Routledge
- [45] Hunt A, Kirk R (2000) Mapping strategies for musical performance. *Trends in Gestural Control of Music* pp 231–258
- [46] Jayaram S, Vance J, Gadh R, Jayaram U, Srinivasan H (2001) Assessment of vr technology and its applications to engineering problems. *J Comput Inf Sci Eng* 1(1):72–83
- [47] Johnson D, Tzanetakis G (2017) *Vrmin: Using mixed reality to augment the theremin formusical tutoring*. In: *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME-17)*, Copenhagen, vol 33
- [48] Juslin PN (1997) Emotional communication in music performance: A functionalist perspective and some data. *Music perception* 14(4):383–418
- [49] Lanier J (1993) *The sound of one hand*. Whole Earth Review
- [50] Larsson P, Västfjäll D, Kleiner M (2001) The actor-observer effect in virtual reality presentations. *CyberPsychology & Behavior* 4(2):239–246
- [51] Lee SW, Freeman J, Colella A, Yao S, Van Troyer A (2011) Collaborative musical improvisation in a laptop ensemble with lolc. In: *Proceedings of the 8th ACM conference on Creativity and cognition*, ACM, New York, NY, USA, C&C '11, pp 361–362, DOI 10.1145/2069618.2069696, URL <http://doi.acm.org/10.1145/2069618.2069696>
- [52] Lele A (2013) Virtual reality and its military utility. *Journal of Ambient Intelligence and Humanized Computing* 4(1):17–26
- [53] Lyon E, Caulkins T, Blount D, Ico Bukvic I, Nichols C, Roan M, Upthegrove T (2016) Genesis of the cube: The design and deployment of an hdla-based performance and research facility. *Computer Music Journal* 40(4):62–78
- [54] Mackinlay JD, Card SK, Robertson GG (1990) Rapid controlled movement through a virtual 3d workspace. In: *ACM SIGGRAPH computer graphics*, ACM, vol 24, pp 171–176
- [55] Magnusson T (2010) Designing constraints: Composing and performing with digital musical systems. *Computer Music Journal* 34(4):62–73
- [56] Mäki-Patola T, Laitinen J, Kanerva A, Takala T (2005) Experiments with virtual reality instruments. In: *Proceedings of the 2005 International Conference on New Interfaces for Musical Expression (NIME05)*, Vancouver, BC, Canada
- [57] Malpas J (2011) Place and the problem of landscape. *The place of landscape: Concepts, contexts, studies* pp 3–26
- [58] McDermott J, Gifford T, Bouwer A, Wagy M (2013) Should music interaction be easy? In: *Music and human-computer interaction*, Springer, pp 29–47
- [59] McKinney J, Butterworth P (2009) *The Cambridge Introduction to Scenography*. Cambridge University Press
- [60] McMahan A (2003) *Immersion, engagement, and presence: A new method for analyzing 3-D video games*, Routledge (New York), pp 67–86

- [61] McPherson AP, Zappi V (2015) Exposing the scaffolding of digital instruments with hardware-software feedback loops. In: NIME, pp 162–167
- [62] McVeigh-Schultz J, Márquez Segura E, Merrill N, Isbister K (2018) What's it mean to "be social" in vr? mapping the social vr design ecology. In: Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems, pp 289–294
- [63] Mendes D, Caputo FM, Giachetti A, Ferreira A, Jorge J (2019) A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments. In: Computer Graphics Forum, Wiley Online Library, vol 38, pp 21–45
- [64] Merritt T, Kow W, Ng C, McGee K, Wyse L (2010) Who makes what sound?: supporting real-time musical improvisations of electroacoustic ensembles. In: Proceedings of the 22nd conference of the computer-human interaction special interest group of Australia on computer-human interaction, ACM, pp 112–119
- [65] Milgram P, Kishino F (1994) A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77(12):1321–1329
- [66] Mulder AG (1998) Design of virtual three-dimensional instruments for sound control. PhD thesis, Simon Fraser University, Canada
- [67] Nilsson NC, Serafin S, Nordahl R (2014) Establishing the range of perceptually natural visual walking speeds for virtual walking-in-place locomotion. *IEEE transactions on visualization and computer graphics* 20(4):569–578
- [68] Norman DA (1999) Affordance, conventions, and design. *interactions* 6(3):38–43
- [69] Nowak KL, Biocca F (2003) The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 12(5):481–494
- [70] Pinho MS, Bowman DA, Freitas CM (2002) Cooperative object manipulation in immersive virtual environments: framework and techniques. In: Proceedings of the ACM symposium on Virtual reality software and technology, ACM, New York, NY, USA, VRST '02, pp 171–178, DOI 10.1145/585740.585769, URL <http://doi.acm.org/10.1145/585740.585769>
- [71] Poepel C (2005) On interface expressivity: A player based study. In: NIME, Citeseer, vol 5, pp 228–231
- [72] Prpa M, Schiphorst T, Tatar K, Pasquier P (2018) Respire: a breath away from the experience in virtual environment. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, ACM, p Art10
- [73] Razzaque S, Swapp D, Slater M, Whitton MC, Steed A (2002) Redirected walking in place. In: EGVE, vol 2, pp 123–130
- [74] Reeves S, Benford S, O'Malley C, Fraser M (2005) Designing the spectator experience. In: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM, New York, NY, USA, CHI '05, pp 741–750, DOI 10.1145/1054972.1055074, URL <http://doi.acm.org/10.1145/1054972.1055074>
- [75] Rose T, Nam CS, Chen KB (2018) Immersion of virtual reality for rehabilitation-review. *Applied ergonomics* 69:153–161

- [76] Rossmly B, Wiethoff A (2019) The modular backward evolution-why to use outdated technologies. In: NIME, pp 343–348
- [77] Sanchez-Vives MV, Slater M (2005) From presence to consciousness through virtual reality. *Nature Reviews Neuroscience* 6:332–339
- [78] Serafin S, Erkut C, Kojs J, Nilsson N, Nordahl R (2016) Virtual reality musical instruments: State of the art, design principles, and future directions. *Computer Music Journal* 40(3):22–40, DOI 10.1162/COMJ_a_00372
- [79] Slater M (2009) Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1535):3549–3557
- [80] Slater M, Steed A (2000) A virtual presence counter. *Presence* 9(5):413–434
- [81] Slater M, Wilbur S (1997) A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 6(6):603–616
- [82] Slater M, Usoh M, Steed A (1995) Taking steps: The influence of a walking technique on presence in virtual reality. *ACM Trans Comput-Hum Interact* 2(3):201–219, DOI 10.1145/210079.210084, URL <http://doi.acm.org/10.1145/210079.210084>
- [83] Slater M, Usoh M, Steed A (1995) Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI)* 2(3):201–219
- [84] Slater M, Linakis V, Usoh M, Kooper R, Street G (1996) Immersion, presence, and performance in virtual environments: An experiment with tri-dimensional chess. In: *ACM Virtual Reality Software and Technology (VRST 1996)*, pp 163–172
- [85] Slater M, Pertaub DP, Steed A (1999) Public speaking in virtual reality: Facing an audience of avatars. *IEEE Computer Graphics and Applications* 19(2):6–9
- [86] Slater M, Guger C, Edlinger G, Leeb R, Pfurtscheller G, Antley A, Garau M, Brogni A, Friedman D (2006) Analysis of physiological responses to a social situation in an immersive virtual environment. *Presence* 15(5):553–569, URL <http://dblp.uni-trier.de/db/journals/presence/presence15.html#SlaterGELPAGBF06>
- [87] Strick M, de Bruin HL, de Ruiter LC, Jonkers W (2015) Striking the right chord: Moving music increases psychological transportation and behavioral intentions. *Journal of Experimental Psychology: Applied* 21(1):57
- [88] Swift B (2013) Chasing a feeling: Experience in computer supported jamming. In: *Music and human-computer interaction*, Springer, pp 85–99
- [89] Trueman D (2007) Why a laptop orchestra? *Organised Sound* 12(2):171–179
- [90] Vincs K, McCormick J (2010) Touching space: Using motion capture and stereo projection to create a “virtual haptics” of dance. *Leonardo* 43(4):359–366
- [91] Wanderley MM, Depalle P (2004) Gestural control of sound synthesis. In: *Proceedings of the IEEE*, pp 632–644

- [92] Wu Y, Zhang L, Bryan-Kinns N, Barthelet M (2017) Open symphony: Creative participation for audiences of live music performances. *IEEE MultiMedia* 24(1):48–62
- [93] Zappi V, McPherson A (2018) Hackable instruments: Supporting appropriation and modification in digital musical interaction. *Frontiers in ICT* 5:26
- [94] Zappi V, Brogni A, Caldwell DG (2010) Osc virtual controller. In: *NIME*, vol 10
- [95] Zappi V, Mazzanti D, Brogni A, Caldwell DG (2011) Design and evaluation of a hybrid reality performance. In: *NIME*, vol 11, pp 355–360
- [96] Zappi V, Mazzanti D, Brogni A, Caldwell D (2012) Concatenative synthesis unit navigation and dynamic rearrangement in vrgrains. In: *9th Sound and Music Computing Conference*
- [97] Zappi V, Pistillo A, Calinon S, Brogni A, Caldwell D (2012) Music expression with a robot manipulator used as a bidirectional tangible interface. *EURASIP Journal on Audio, Speech, and Music Processing* 2012(1):1–11
- [98] Zhai P (1998) *Get real: A philosophical adventure in virtual reality*. Rowman & Littlefield