A REVIEW OF USER ROLES IN COLLABORATIVE VR MUSIC-MAKING ENVIRONMENTS: TOWARDS ROLE DESIGN GUIDELINES WITH *WAM JAM PARTY* AS A CASE STUDY

Adam Mir-sadjadi, Michel Buffa, Marco Winckler and Florent Robert
Université Côte d'Azur, France
firstname.name@univ-cotedazur.fr

ABSTRACT

Collaborative music making (CMM) in virtual reality (VR) enables embodied, spatial, and socially situated forms of musical interaction. However, when multiple users manipulate shared musical artifacts (e.g., patches, timelines, or parameter sets), concurrent actions can lead to interruptions and conflicting edits that disrupt creative flow.

This paper reviews how user roles are addressed in collaborative VR music-making systems. We examine roles as musical or social participation (e.g., performer, tutor, audience) and as authority over shared artifacts, determining who can modify what and under which conditions. We identify recurring coordination mechanisms such as personal territories, awareness cues and rule-based interaction constraints, as well as design tensions including symmetry vs. expertise and freedom vs. protection.

These observations lead to a set of design guidelines that we examine through the case study of *WAM Jam Party*, a node-based collaborative VR patching environment based on Web Audio Modules (WAMs). We discuss how configurable roles and permission schemes could support collaborative composition while preserving creative flow in immersive music-making environments.

Keywords: collaborative music making, virtual reality, user roles, permissions, conflicts, shared virtual environments, computer supported collaborative work, web audio modules

1. INTRODUCTION

Context and motivation: Overlapping modifications made by different users are a well-known challenge in computer-supported collaborative work (CSCW), particularly in shared editing systems where concurrent actions may lead to conflicts or inconsistent states [16, 20]. Similar issues appear in Collaborative Music Making (CMM) systems, where multiple participants manipulate shared musical structures or parameters in real time. In such contexts, conflicting actions can create friction in the creative process or produce unintended sonic results that disrupt group flow. The reacTable, for instance, illustrates how simultaneous manipulation of a tangible shared interface can lead to conflicting user intentions during performance [21].

In virtual environments, these issues take the form of concurrent edits that must be either managed at the system level or prevented through interaction design. Specifically, Virtual Reality (VR) introduces Shared Virtual Environments (SVEs) where users can meet up and interact through embodied and continuous interactions as opposed to traditional media which rely on discrete event-based input (e.g., mouse clicks). While these enable strong social presence and shared performance, this real-time gestural interaction allows multiple users to act on the same object at the same time and create friction, making temporal and spatial overlaps more likely and increasing the risk of interference or conflicting modifications between users.

Furthermore, reflections around new audience experiences in VR and the musical metaverse (the collection of music-related SVEs in VR) seem to go in the direction of giving more agency to the spectators and to include them in the experience not only as a passive audience but also as an active part of the experience. Boem et al. frame such VR music experiences as "*a new form of making art, which subvert the conventional roles of the artist fully in charge of the creative process and of a passive audience*" [4] while Bradley still points out that considering the number of HMD owners, some if not most of "*users are likely to attend "at-keyboard" - as viewers, not participants –, watching others experience virtual reality rather than entering it themselves.*" [5].

In this context, user roles and permissions become primary design levers to reduce friction while preserving experimentation, play and inclusive participation in various use cases.

Objectives: This study aims to **(i)** identify user roles in collaborative musical VR experiences, **(ii)** analyze roles, permissions and potential situations requiring explicit design choices, and **(iii)** derive design implications for structuring role distributions in collaborative VR music systems. More broadly, the goal is to explore how roles and permission structures can be designed in VR CMM systems, with *WAM Jam Party* serving as a preliminary testbed application context to reflect upon.

Scope and research questions: To address these objectives, we review prior work across two complementary domains: (a) collaborative VR music instruments and environments, and (b) collaborative music and HCI systems that address role management through mechanisms such as privacy and awareness, territoriality, asymmetric permissions, and coordination under latency. Based on this review, the study addresses the following research questions:

RQ1. What role concepts recur in VR CMM systems?

RQ2. Which coordination and conflict resolution mechanisms (e.g., personal space, locking, turn-taking, moderation) are reported, and how do they support creative flow while preserving collaboration?

RQ3. How can these be translated into configurable role distributions and design guidelines across different CMM contexts (free jam, lesson, performance, game)?

Here, *free jam* denotes unstructured co-creation, *lesson* denotes teacher-led practice with explicit authority, *performance* denotes audience-facing sessions with show control, and *game* denotes rule-driven collaboration or competition.

We aim to account for different types of CMM environments and derive guidelines applicable for a broad range of systems, while focusing our analysis primarily on applications similar to *WAM Jam Party*.

Target system: Inspired by *Sequencer Party* [12], *WAM Jam Party* [11] is a collaborative VR immersive environment in which multiple users assemble modular audio systems in real time by connecting Web Audio Modules (WAM) plugins [10] (note generators, instruments, effects) into interactive audio graphs with spatialized outputs. Users interact through head-mounted displays (HMDs) and handheld controllers thanks to which they can grab, move, connect and edit nodes in 3D. Some modules expose custom interfaces like a piano-roll editor for note sequencing or embodied virtual reality musical instruments (VRMIs, VR-specific VMIs) such as a virtual drum kit [25].

All participants share the same musical environment: plugin states, parameter values, musical patterns, tempo and spatial configurations are synchronized several times per second. However, the current prototype does not implement explicit roles, locking, or ownership mechanisms. As a result, several users may simultaneously manipulate the same musical artifact. While this openness encourages experimentation and improvisation, it can also disrupt collaborative flow, for example when the piano-roll editor records drum patterns played by one user while another participant edits the same sequence.

An open-source WebXR prototype is publicly available [1] and actively maintained [2], providing a concrete testbed for experimenting with role distributions and permission mechanisms in collaborative VR music systems.

Contributions: The proposed study presents a structured review of user roles in collaborative VR music-making environments and derives design guidelines for CMM SVEs. It further explores these insights through an exploratory case study of *WAM Jam Party*, proposing an adjustable role-based permission model intended to mitigate coordination challenges, supported by an open-source prototype.

This paper is structured as follows. Section 2 describes the materials and methods used to construct and analyze the literature corpus. Section 3 presents the results of the state-of-the-art review and exposes guidelines derived from its synthesis. Section 4 investigates these guidelines implications for *WAM Jam Party* as a case study. Section 5 discusses various aspects of this work. Finally, Section 6 concludes and outlines directions for future work.

[1] https://wamjamparty.i3s.univ-cotedazur.fr/

[2] https://github.com/doriangirard9/musical-multiverse-vr

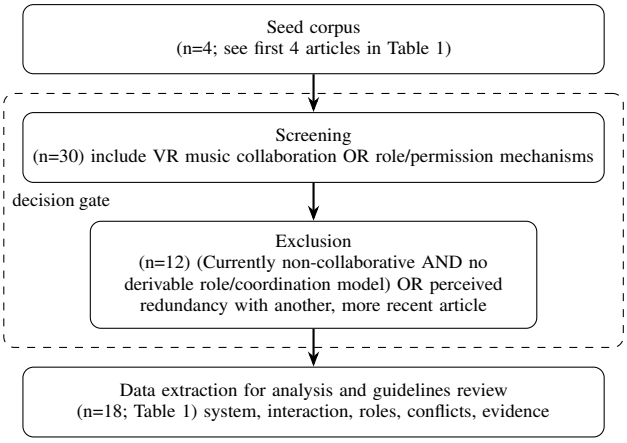

Figure 1. Review pipeline used for this paper selection.

2. CORPUS CONSTRUCTION, DATA EXTRACTION AND SELECTION FLOW

The corpus was constructed through a quasi-systematic review, combining structured search and screening steps, without following the full protocol of a formal systematic review. We started from 4 seed articles (see the first 4 articles in Table 1) then expanded the corpus using three complementary strategies: Keyword search on Google Scholar [3], backward and forward snowballing. We included papers explicitly discussing collaborative VR music-making (shared VRMIs or spaces), role asymmetries or permission structures, privacy/awareness or territoriality mechanisms, and VR-specific constraints (embodied interaction, spatial audio, social presence, spectatorship).

Additionally, some references have been added to the corpus during the reviewing process to get a more exhaustive overview by including relevant industry products (PatchXR, Virtuoso) [26, 27]. We excluded papers that are non-collaborative, purely technical without interaction design implications, or do not provide interpretable evidence about roles or coordination. For each paper we extracted, when possible: system type, interaction modality, explicit and/or implicit role definitions, conflict/coordination mechanisms, and reported benefits or limitations relevant to creative flow. The process is summarized in Figure 1.

3. STATE OF THE ART REVIEW: RESULTS AND SYNTHESIS

3.1. Corpus characterization

Table 1 summarizes the core and extended works analyzed in order to conduct this review and further provides topic relevance comments as well as a comparative overview of key systems and industry products discussed.

The corpus spans 2002–2025, mixing VR-native CMM systems with adjacent CSCW and non-VR music work

[3] The final string built was:
`("multi-user" OR "multiplayer" OR "multiple users" OR "collaboration") AND ("role" OR "can modify" OR "can alter" OR "authorizations" OR "privileges") AND ("music" OR "musical" OR ("interactive" AND "sound" AND "experience")) AND ("VR" OR "virtual reality" OR "virtual installation")`

Paper/ Ref	Year	Venue / Editor	Type	VR?	System	Platform - Technology (Availability - Pricing)	Short description / Context	Role focus / Relevance to the topic
[1]	2025	ACM IMX	Conf. paper	Yes	Various XR installations	-	45 public VR installations review	Designers-oriented review, role taxonomy including organizational roles
[3]	2024	NIME	Conf. paper	Yes	Custom VRMIs	HMDs (Web) - NAF, PdXR (Open source / Free)	Shared VRMIs ; paired user study	Symmetric vs. asymmetric roles, parameter-split control
[28]	2024	NIME	Conf. paper	Yes	*Gamelan Land*	Meta Quest - Unity (N/A - N/A)	Multiplayer VR game; 14 users	Role structure from traditional music communities, management roles
[9]	2025	ISCC	Conf. paper	No	*Sequencer Party*	PC (Web) - WAMs (Online / Free)	WAM-based desktop CMM	Real-time sync. constraints, sessions sharing, WAM-oriented workflow
[24]	2021	Audio Mostly	Conf. paper	Yes	*LeMo*	HTC Vive - Unity (Open source / Free)	CMM SVE, 26 users	Personal and public territories, awareness, progressive boundaries
[18]	2010	NIME	Article	No	Custom CMM Soft.	PC - Java, SuperCollider (N/A - N/A)	Co-located Desktop CMM	Privacy/awareness, empirical support for private spaces before publishing
[19]	2025	Springer	Book chapter	Yes	*Sound Blocks VR*	Meta Quest - Unity (Open source / Free)	Accessible hand-tracked VRMI	Accessibility for novice musicians, role simplification for non-expert users
[7]	2012	Behav. & IT	Journal	No	*Daisyphone*	PC - Java (Online but no longer hosted / Free)	Collaborative creativity study	Mutual engagement and group flow theory informing role/awareness design
[8]	2004	NIME	Article	No	*Daisyphone*	PC - Java (Online but no longer hosted / Free)	Early remote desktop CMM	Shared control, early evidence for remote musical coordination challenges
[20]	2002	CSCW	Journal	No	Various groupwares	-	Real-Time Groupwares	Workspace awareness framework, insights for cues in shared editing
[13]	2016	NIME	Article	No	*Monad*	PC - C++, OpenFrameworks (Open source / Free)	Networked virtual environment	Game mechanics in a CMM context: role structuring via constraints, scores
[14]	2020	NIME	Article	Yes	*INVISO*	PC & HMDs (Web) - three.js, Web Audio (Open source / Free)	Cross-platform sonic design tool	Cross-platform collaboration, supports role differentiation across devices
[4]	2023	IS2	Conf. paper	Yes	*Musical Metav. Playgrounds*	HMDs (Web) - NAF, Tone.js, Essentia.js (Open source / Free)	Web-based CMM SVE	Browser-based shared creation spaces, audience involvement
[15]	2023	JAES	Journal	Yes	*Orchestra* (Toolbox)	HMDs (Web) - NAF, Resonance Audio, PdXR (Open source / Free)	Web-based VR performance toolbox	Scenic performance roles, operational separation in live music contexts
[2]	2003	NIME	Article	No	Various CMM systems	-	Early CMM Systems overview	Restricted control for accessibility, motivates simplified roles for novices
[17]	2020	iJET	Article	No	Custom application	Android - N/A (N/A - N/A)	Music learning Gamelan app.	Guided learning for novices, virtual instructor role fulfilled by the system
[26]	2022	PatchXR AG	Software	Yes	*Patchworld*	HMDs, PC - N/A (Online stores / Paid: 29.99$)	Industry VR CMM SVE	Musical experiences with musical roles, world ownership and permissions
[27]	2022	Fast Travel Games	Software	Yes	*Virtuoso*	HMDs, PC - N/A (Online stores / Paid: 19.99$)	VR instrument performance environment	Industry VR music-making environment; supports contribution-based collaboration

Table 1. Core and extended corpus for role/permission design in collaborative VR music systems

providing transferable role design mechanisms. Approximately half focus on immersive VR environments or shared virtual instruments [1, 3, 28]. The remainder address coordination challenges persistent in VR contexts [2, 7, 20] such as awareness, ownership or concurrent editing. The systems cover shared virtual instruments, networked music environments, browser-based patching, performance exhibits, and educational applications, with a mix of user studies [3, 24], SVEs or toolkit papers [14, 15] as well as literature reviews [1, 2], revealing recurring patterns in role distribution in both early collaborative music interfaces and recent immersive VR environments.

3.2. Role characterization in VR CMM systems

Across the reviewed corpus, two complementary perspectives emerge for describing user roles in collaborative VR music systems. 1) A user role can be understood as a **musical/social role** (what the person is *there to do*). Examples include performer, conductor or bandleader, tutor, learner, active or passive spectator, or organizer. 2) It can also be understood as an **authority/interaction role** (what the person is *allowed to change*). Examples include owner (full control), editor (can modify shared artifacts), contributor (can add but not overwrite), observer (read-only), or moderator (can modify other users' access rights).

These roles are closely linked to permission structures (e.g., adding, deleting or rewiring nodes, managing access), their scope (own, shared, or global), and to the coordination mechanisms users mobilize to synchronize or communicate with each other (e.g., cues, feedback). This distinction helps avoid a common pitfall: a "performer" or "conductor" may still require restricted authority, while a "learner" may need elevated permissions depending on the context. In many cases, identifying these aspects in prior work requires examining experimental setups and the types of control granted to users, particularly when roles and permissions are not explicitly described.

The following section examines how role configurations appear across different collaborative contexts.

3.3. Multi-role perspectives in different use contexts

Several use cases appear across the reviewed works. While some contexts may overlap depending on the situation (e.g., a user may learn during a free jam), we identify context-specific implications that should be considered when approaching multi-role design in CMM SVEs.

3.3.1. Learning, lessons and accessibility

The benefits of CMM in virtual environments for supporting music learning (both theory and practice) and social interaction have been widely studied, including in industry platforms such as *Patchworld* or *Fortnite* [6]. Fanani et al.'s work on a rule-based learning application showed such

benefits by reporting an evaluation made with two experienced musicians recruited to train two novice groups. The first group of users played inside a virtual instructor-led music game between training sessions and achieved "*excellent skill scores*" faster than the second control group who didn't [17]. Collaborative aspects in CMM applications can set the ground for new kinds of lessons inside virtual environments where the role of the teacher impersonated by a user could emerge.

Early work on collaborative musical interfaces emphasizes that technology should primarily support social interaction rather than being the focus of the experience [2]. Many systems therefore simplify musical control so that novice users can participate without needing to master structures such as scales or harmony. This strategy is particularly relevant in public installations and walk-up-and-play scenarios, where interaction must be quickly accessible.

Sound Blocks VR similarly foregrounds accessibility considering motor impairments in VR instrument design, presenting an Accessible Digital Musical Instrument (ADMI) that is "*also suitable for many other applications, including music education in young learners and new forms of expressivity*" [19]. This collaborative instrument is fully compatible with controllers as well as hand-tracking (preferred by motor-impaired users) and allows for the use of "*presets for musical scales*" and advanced parameter under a button menu. Thanks to this instrument design, an "*impaired musician can participate and even become the soloist at the center of the performance*", revealing that entire musical roles may be inaccessible to some users depending on the related instruments designs. Identifying accessibility barriers in role and interaction designs should remain a top priority to include the broadest audience possible while enabling collaboration and expressivity.

The takeaway from this first use context of CMMs is that a primary aspect in role design should be to ensure accessibility, accommodate novice or expert profiles of players and consider emerging roles such as a teacher role for lessons-oriented applications. Several role-related game mechanisms can also improve the learning context and are addressed in section 3.3.3.

3.3.2. Scenic and organizational roles in performances and exhibits contexts

Azevedo et al. reviewed 45 works on immersive public installations and conducted a workshop with designers and event organizers [1]. They identified several recurring roles in specific frames inside and around the installations in their synthesis. Their analysis found several layers of participation surrounding immersive installations, ranging from peripheral spectators and queued participants to operational staff (operators, technicians) and fully immersed participants or performers. This shows that role design must extend beyond end-users to include facilitators and operational staff when there might be one. Additionally, for exhibits happening fully inside VR experiences (imagining several stages inside a shared space), many roles described here in the real world could be ported to the virtual space inside of HMDs (e.g., with passersby walking through the different installations), making this taxonomy a precious resource to identify emerging roles in SVEs.

In parallel, *Orchestra* [15] provides a toolbox for live audiovisual performances in web-based metaverse contexts based on AFrame/Networked-AFrame [4] (NAF) that might be approached as conventional scenic experiences (i.e. with a stage or several installations, lighting, effects...), furthermore setting the ground for operational roles such as scene management and show control. The toolbox integrates a sound broadcast component (e.g., speech/applause), PdXR [5] components with which it is "*possible to perform together on virtual instruments or to create interactive virtual installations*", and live-coding components integrating three Esolang live-coding languages (languages with a reduced set of instructions for "*conceptual, creative, and sometimes artistic [...] designs*" rather than general-purpose use) enabling live-coding in VR through virtual keyboards thanks to their simplified design: Bytebeat, StackBeat, and IBNIZ.

Collaboration examples in the article only show two users performing on live coding terminals but they don't seem to be modifying the same source code in the figure shown (the article mentions the source code of components is shared amongst all users but no details are provided about concurrent code modifications). Although user roles aren't explicitly discussed, this work underlines a variety of use cases that might require specific responsibilities to fit in a performance, giving further relevance to a role distribution for performance-oriented CMMs following a detailed structure as the one from Azevedo et al.'s study.

Including the audience inside performances can also be achieved with new interactions designed for audience-related roles. Blaine and Fels highlighted several instruments as early as 2003 which fostered a sense of community by encouraging movements such as dancing with strangers [2]. SVEs offer new meeting spaces for dematerialized performances which could include such interactions to act upon the music produced by performers or add some interactive aspects to usually passive audience roles.

These observations highlight that performance-oriented CMM systems can involve an equal if not broader range of roles than traditional musical ensembles. In addition to performers, roles such as operators, technicians or facilitators may contribute to managing the experience and assisting participants. Role design therefore extends beyond musical interaction to include organizational and operational staff.

3.3.3. Free jam / Game mechanics: reward system, role-play & virtual roles

Cakmak et al. introduce Monad, a collaborative music system that integrates game mechanics through an internal point economy [13]. Players must spend points to perform actions, while contributions can be rewarded either by other players or automatically by the server - which users pre-

[4] a web multi-user VR framework

[5] `https://github.com/vroom-space/pdXR`

ferred as it didn't require them to interrupt their work to reward other players. Their evaluation shows that structured constraints can stabilize collaboration but may also influence the musical outcome and interaction dynamics.

Syukur et al. document the role hierarchy of Gamelan (Indonesian music ensembles) studios - owners, managers, instructors, musicians, singers, visitors and clients - in which the total number of instrument ranges from 30-60 and above. This hierarchy was translated into a multiplayer VR game designed for users to play on a long-term basis called *Gamelan Land* [28]. In the experiment, active participants were limited to the instructor, musician, singer and audience while the owner and manager's roles are simplified to scheduling practice and paying studio members and performed by the system like in *Monad*'s second setting. Similarly, the story's narration revolves around the activity of collecting points. Although the game avoids conflicts by assigning one instrument and role per user, the real-world hierarchy shows that traditional ensembles already rely on structured role distributions, making role design a part of cultural preservation as well as a lever to reduce friction.

In addition to the actual user roles, non-playable characters (NPC) virtual roles can also enhance the user experience, for instance by accelerating novice players' onboarding through tutorials, whose importance has been pointed out by several studies reported in Azevedo et al.'s work [1]. This insight is highlighted in Fanani et al.'s work where a virtual instructor is implemented inside a Gamelan learning game to teach players how to play Gamelan instruments [17]. This role is fulfilled by the system which runs a model adapting to the player's level to give him or her feedback depending on the system's expectations. Considering Nintendo *Wii Fit* virtual instructor's popularity as of today, virtual roles should be accounted for to facilitate users' immersion and to fulfill specific goals as they can provide the drive needed to partake in the experience, or constitute an obstacle if not deemed convincing enough.

While these examples highlight how roles emerge from different collaboration contexts, they do not fully explain how access to shared musical artifacts is technically regulated. The following section therefore examines mechanisms used to structure authority, access and coordination.

3.4. Role-based access control (RBAC) as a conflict prevention and experience design lever

3.4.1. Symmetric vs asymmetric roles in shared VMIs

Boem and Turchet introduce the *Musical Metaverse Playgrounds*, a web-based VR application hosting shared virtual sonic experiences [4]. Two "*playgrounds*" are presented. In both of them, users can move through the environment, communicate through voice speech and interact with 3D objects in the scene through gestures. In the first playground, users can generate 3D synthesizers and play them simultaneously, while the second focuses on multimodal interactions by enabling sound generation accounting for the users microphone inputs which is visually represented through a series of animated vertical bars. Additional visualizations which can only be controlled by their creators "*to avoid confusion*" can be created by the users. 8 users were paired with an examinator to conduct experiments in both playgrounds through a thinking-aloud protocol that was followed by an unstructured interview. These experiments give interesting insights considering roles and permissions as perceived by the users. Notably, in the first playground, three participants "*found confusing that any connected user could modify the sound generators, even if they weren't the creators*", pointing out that "*such type of collaborative activity (i.e., changing together the parameters of a sound generator) is dependent on the type of community involved and on how much a person will trust the other users*". Consequently, "*They highlighted that it should be important to set some boundaries*" and suggested the implementation of a mechanism to "*share the ownership of the generators they have created*".

Shared VMIs also raise a fundamental shared-control problem mentioned in this paper's introduction: performers can simultaneously modify parameters of a single interface [3]. Boem et al. compare three VRMIs with distinct role configurations - strongly asymmetric (a player moves the sound source while the other changes the sound it produces), partly asymmetric (identical controls but managing different synthesis parameters of the instrument), and fully symmetric (both users can do the same things) - in a user study with 12 participants. The strongly asymmetric instrument was perceived as confusing ("*I do not quite understand what my role is*"), partly because one player was in charge of moving the sound source, making the other player unable to hear the spatialized feedback of what he was doing. By contrast, the partly asymmetric design achieved the best usability (SUS) and creativity-support (CSI) scores, and also fostered stronger social presence, suggesting a link between role clarity and collaboration quality. It is also notable that the fully symmetric instrument incited players to coordinate between themselves ("*You do the kick drum, I do the snare*"), preventing conflicting intentions.

As seen in learning contexts, educational VR music systems might also require asymmetric control to support instruction and assessment or to allow novice and expert players to play together without creating chaos. Boem et al.'s conclusion over the three types of instruments introduced also go this way, indicating the strongly asymmetric design "*may be better suited for more experienced players. Conversely, systems like those of [the 2 other designs] appear to be more ideal for novices*". Rule-based learning applications show how structured feedback and constrained interaction can guide novice learning, which translates into teacher-led authority models in VR lessons [17]. A user impersonating the teacher could need to be allowed greater control over the session and manage access for a single or several students to specific sets of nodes or parameters depending on the size of the "class".

In gaming contexts, asymmetry should not introduce imbalance amongst players. *Monad* players participated in an evaluation combining local and remote players. When discussing the possibility of adding embodied interactions to local players in the physical space, said players "*expressed*

that physical gestures would go against the democratic nature of the performance by way of overpowering the role of the remote players in the audiovisual output" [13].

3.4.2. Cross-platform role design and parameter-split

INVISO [14] introduces a cross-platform interface supporting VR and AR for creating virtual sonic environments on desktop or inside of HMDs. While collaboration is mentioned as a future perspective of the work presented, the cross-platform aspect is interesting to account for, as users might access the SVE from different devices depending on their social role and might simultaneously control different parameters of shared audio sources in the scene.

Boem et al.'s study [3] tested such a parameter-split instrument, where one user could move the sound location while the other user changed the sound it emitted. Their study tends to disavow the viability of this solution as one of the users was not able to hear what he was doing due to the other player's manipulations of the localization. We might still wonder if a different design would have made this solution viable (by allowing the user changing the sound to hear his work in a private space for instance). At the same time, they report splitting control over different sound synthesis parameters of the instrument is well perceived by users, emphasizing the potential of sharing controls if done thoughtfully.

The cross-platform aspects also resonates with Azevedo et al.'s categorization of the Technician role [1], who monitors the VR experience from a desktop environment. As part of the scenic experience, a captation of performances could be done by technicians on desktop in real-time through a set of cameras placed in advance in the scene or controlled by spectators. The output could be displayed on a flat-surface screen to enable a broader audience outside of HMDs to attend the experience, as well as opening the possibility for some users to contribute from an actual keyboard, enabling the use of traditional live coding languages in addition to Esolangs for toolboxes such as Orchestra [15], or the appearance of additional roles compatible from desktop access such as moderators.

3.4.3. Session sharing in current CMM spaces

Patchworld [26] and *Virtuoso* [27] are the two main industry references for VR CMM SVEs while *Sequencer Party* [9] is a node-based, web-based CMM environment and the "ancestor" of *WAM Jam Party*. All three enable collaboration between users and implement a declination of role-based access control. Designed as fully functional solutions operating since several years, insight can be gained by analyzing their permission models.

In *PatchWorld*, users can create public or private "*worlds*" that they own [6]. Public worlds can be seen by anyone. Private worlds can be seen if shared with at least the "View" permission level, or if they are published. Both private and public worlds can be shared with 2 additional permission levels: Edit (a designated user can edit) or Open for Edit (anyone can edit). Worlds can also "*be linked together to create interconnected experiences*". Additionally, during live sessions, the host (creator) can access a Mic/Touch/Edit permission menu listing each users permissions on his wrist [7], allowing for fine-grained control.

Similarly, in *Sequencer Party*, *organizers* can create, edit or delete sessions of their own in which they have every right to add, change or remove tracks and assets (nodes). They can invite guests (passive audience) or collaborators to their sessions. Collaborators are either **performers** (who *"can manipulate parameters of WAM components but cannot add or delete such components"*) or **editors** (who can *"add, delete or rewire component; change plugins and edit generative logic or sequence content"*).

Virtuoso, however, does not currently support real-time collaboration over internet. Users can share their musics in a Community Library, where other users can pick them up and contribute to make "asynchronous" collaborative songs. All players participating are then listed in the song as contributors [8]. Still, local collaboration is supported through *Ableton Live*, mitigating long-distance delay issues. In such a context, the first user to connect determines the tempo. Any connected user can alter it, and races can occur if several users do it simultaneously. Users cannot be prevented from modifying the tempo if they want to [9]. Additionally, extreme tempo changes in opposing direction can make "*the beat timeline diverge - However, Link will always keep the phase of the apps in sync*". The recording's responsibility can be split by having each participant record their session for later merging, or delegated to one user by mixing everyone through a single mixer and capturing the stereo output or the multi-track.

These systems illustrate how role-based access control can structure collaboration in practice but permissions alone are not sufficient to support smooth collaboration. Participants must also remain aware of each other's actions and coordinate their activities in real time. The next section therefore examines coordination mechanisms supporting mutual engagement and group flow.

3.5. Coordination mechanisms regarding mutual engagement and group flow

3.5.1. Awareness

Bryan-Kinns and Hamilton argue "*rich, creative collaboration*" depends on whether participants experience *mutual engagement*, a state of group flow in which users respond to each other's actions and ideas [7]. Prior CSCW works shows that explicit locking mechanisms become less necessary when collaborators are sufficiently aware of each other's actions [20]. Building on this ground, they argue the key features supporting mutual engagement are: **(i) Mutual awareness of actions:** Highlighting new contributions, identifying who contributed, where..., **(ii) Shared**

[6] https://wiki.patchxr.io/patching/worlds

[7] https://wiki.patchxr.io/general/live-sessions

[8] https://virtuoso-vr.com/faq

[9] https://www.ableton.com/fr/live-manual/11/synchronizing-with-link-tempo-follower-and-midi/

and consistent representations: Views on the joint product should be the consistent across users, **(iii) Mutual modifiability:** Being able to modify each other's contributions - Implying an egalitarian approach to role assignment rather than role enforcement., **(iv) Annotation:** Being able to converse in and around the shared product.

As a result of their study on the *Daisyphone*, a 2004 shared VMI applying these features, they suggest that "*role assignment emerges naturally and does not need to be explicitly built into the interface*" [8]. Comparing two versions of the *Daisyphone*, the authors report that by simply changing note persistence to a decaying version, sessions became unstructured and resembled conventional music-making in real life a lot more as users had to focus and less time to communicate, showing that role-design (or its absence) is tightly coupled to the experience goal.

Analyzing their results, they finally argue that mutual modification in CMM is *not as strong an indicator of mutual engagement as coherence and musicality of the final product*. Users who felt the most engaged tended to be the ones with less activity, who modified the contributions of others less and rather focused on their own contributions to the shared piece. This was also facilitated by the presence of annotations and IDs which allowed for less time dedicated to organizing the tasks as users could see what others were contributing and collect information quickly. Authors argue accountability through IDs may also have made the users more cautious of their contributions. Earlier studies on the *Daisyphone* also showed the importance of localization (being able to reference parts of the shared product to discuss them with others or propose revisions) and implemented it through an annotation system using the mouse to draw, write or circle parts of the composition. This is supplemented by a history system through automatic annotations when notes are set / unset which allows one to keep track of who made which edit and when, providing context without wasting the restricted communication windows.

The importance of communication is pointed out in several other studies with different modalities, from the *Musical Metaverse Playgrounds* where "*All participants commented positively on the presence of a real-time voice chat system*" [4] to *LeMo* where participants took turns using a virtual pen to write "*I make*" or "*You do*" in the VR space [22] and *Monad* which implemented a chat system to compensate for the absence of visual cues band members can usually give each other in real life performances [13].

Identification of who is contributing has also been a concern for *Monad* with users showing a "*strong preference for a clear representation of each user's presence within the environment*" and "*feeling dissatisfied when their actions were not evident in the audiovisual output*".

3.5.2. Privacy and territories

A key finding is the value of separating exploration from publication. Fencott and Bryan-Kinns show that musicians make extensive use of private workspaces when available to them [18]. They report users defaulting to a divide-and-conquer approach when it's not available, where each player takes on an informal role (e.g., focusing on bassline or drums). Personal spaces rather let users experiment in private before sharing their work for "*selection, scrutiny and further revision*". In a thorough study, Men et al. test four sound attenuation configurations with 52 users composing music in pairs in *LeMo*, a VR CMM SVE [24]. They revealed that adding personal space improves individual creativity, but rigid wall boundaries increase distance between collaborators and reduce group edits. Their key contribution is the concept of acoustic attenuation as a fluid personal space: the system gradually attenuates what a user hears from the shared space, enabling continuous adjustment between full isolation and togetherness. They derive five design implications for SVEs, with three specifically for CMM SVEs: **(i)** They should provide users with personal spaces, **(ii)** Augmented acoustic attenuation is an efficient way to do so for audio-related tasks. Generally, rigid walls should be avoided but are still better than no personal space [23] (a rigid space with mobility is better than a fixed one too), **(iii)** The level of attenuation defines the level of "personalness", therefore it should be configured accordingly and made configurable for users to adjust.

Awareness and privacy are key factors that will define how users approach the concept of roles. When no formal role distribution is given, users will develop strategies to organize themselves whether through turn taking or divide-and-conquer. Designing with awareness and privacy principles in mind can reduce the need for a formal role distribution system as users naturally tend to focus on their own part of the work when provided with personal spaces and clear cues as to what others are doing. Collaboration primarily occurs through communication which can preempt most conflicts (users' actions are less likely to overlap). Implementation of such concepts should be prioritized (knowing it will affect how users approach the experience) in order to identify the remaining conflicts that a formal role system could solve if still needed.

3.6. Cross-paper synthesis: role design guidelines

Across the reviewed works, recurring patterns appear regarding role distribution, coordination mechanisms and conflict management. Instead of converging towards a single role model, the corpus allowed us to derive 6 design guidelines for role design in collaborative CMM SVEs.

G1 – Consider the broader ecosystem of roles: Collaborative VR music systems involve more actors than performers alone. Immersive installations highlight roles such as facilitators, technicians or operators, as well as multiple forms of audience participation ranging from passive spectators to active participants. Educational systems further introduce teacher–learner hierarchy while gaming contexts show virtual roles can complement user roles too. Role design should therefore account for this broader ecosystem surrounding the musical activity.

G2 – Support diverse expertise levels: Many systems aim to accommodate both novice and expert users. Accessible interaction design can support novice participation while preserving expressive potential for experienced musicians. In such contexts, configurable role asymmetry can

help structure collaboration and avoid chaos or errors, for instance in learning or performance preparation scenarios.

G3 – Use explicit roles only when necessary: Explicit role assignment is not always required and can alter the user experience. In several systems, users naturally adopt informal coordination strategies such as turn-taking or task division when roles are not predefined [22]. However, explicit roles may still be useful in contexts where authority must be enforced or to support roleplaying [28].

G4 – Prioritize workspace awareness: Workspace awareness is a key mechanism to prevent conflicts. Participants benefit from cues indicating *who is interacting with what, where, and when*, supported by communication channels such as speech, chat, or visual annotations [20]. Visibility of contributions is expected, as users report frustration when their actions are not clearly identifiable. Roles should be understandable to participants to avoid confusion [3].

G5 – Provide personal territories for experimentation: Musicians frequently prefer to experiment before exposing their work to the group. Personal territories allow the separation of exploration and shared production. In VR environments, augmented acoustic attenuation has been shown to provide a flexible mechanism for modulating privacy between shared and isolated listening conditions [24], letting users adjust their level of "personalness".

G6 – Keep ownership and editing policies configurable: The reviewed systems adopt different approaches to artifact ownership. Some rely on shared editing coordinated socially, while others introduce explicit permission structures. User studies indicate that participants can expect control over artifacts they create [4], suggesting that editing and ownership policies should remain configurable to balance creative freedom with protection from external edits.

Looking back at these guidelines and the review, recurring design tensions appear in collaborative CMM systems, notably *symmetry vs. expertise*, *freedom vs. protection* and *awareness vs. privacy*. The next section explores possible ways to address these tensions through a preliminary reflection on role and permission design in *WAM Jam Party*.

4. CASE STUDY: *WAM JAM PARTY*

This case study is a preliminary exploration of role design in *WAM Jam Party*, proposing mechanisms informed by the literature review to be empirically tested rather than validated solutions. This work is motivated by a continuous development on the app as we are about to implement multi-session persistence to make the application relevant for users to use autonomously beyond being a research subject. Informal testing of *WAM Jam Party* and *Sequencer Party* prior to this work revealed concrete conflicts. For instance, a teacher adjusting volumes and panning to improve his personal listening experience during a class presentation, unaware that these changes were synchronized in the shared session, which confused students as to why their mix kept changing in real-time. In another case, a user recording a drum pattern through a shared piano roll was repeatedly interrupted by another participant simultaneously editing the same interface. These conflicts motivate explicit role mechanisms, though awareness cues and communication means alone may be sufficient as the literature suggested. Making the experience settings customizable by users might appear to be the best solution to allow for various use cases, but the balance between complexity and customizability might be a thin line to walk. In the following subsections, we discuss potential mechanisms and configurations that could mitigate such conflicts by exploring each derived guideline implication for *WAM Jam Party* and proposing an adjustable permission matrix.

4.1. Applying synthesis guidelines to *WAM Jam Party*

G1 – Consider the broader ecosystem of roles: The permission matrix proposed in section 4.3 dissociates scene setup and performance through the technician / performer subdivision. Recordings will be done by technicians once implemented. Considering the node-based nature of *WAM Jam Party*, there is little need for further specification as the type of node used by a performer will determine his temporary musical role (singer, drummer...).

G2 – Support diverse expertise levels: Virtual instructors could be envisioned to support onboarding, but basic tutorials can already preempt chaotic situations. Interactions should be designed to favor accessibility which should benefit every user profile. A specific novice profile might be added to allow for beginners to only use pre-built audio chain parts considering the node-based aspect of *WAM Jam Party*. The pre-built chains will be made accessible for every user as even experts could want to use them.

G3 – Use explicit roles only when necessary: Public sessions will default to free jam with symmetric control, allowing users to self-organize unless explicit roles are required. Still, users should be allowed to save sessions and become owners of their iteration. Sessions will therefore be shareable for iterative improvements through a system similar to *Patchworld*'s or *Sequencer Party*'s. Selecting a context and sharing for edit will provide users with an intermediate-access role while sharing for viewing will invite audience members. Specific roles will be assigned manually by owners and moderators that owners nominate.

G4 – Prioritize workspace awareness: Improvement of avatars (e.g. adding hand animations, name display, gestures) and login implementation will allow for users to easily identify other users in the immersive environment. The addition of communication means such as a voice chat, a 3D pen to point at, circle or write is underway and we also think about adding a text chat for future cross-device participation. Additional cues are considered such as visually linking users to nodes they are modifying (e.g., with colored borders), or displaying notifications indicating where sound modifications originate by highlighting new contributions. Finally, a contribution history should be provided and designed to be usable within HMDs.

G5 – Provide personal territories for experimentation: *WAM Jam Party* already implements an augmented sound attenuation system with multiple spatialized outputs. However, the spatial range of these outputs is currently fixed and should ideally be made configurable. Additional lis-

Context	Role	Add nodes	Edit params	Rewire/ delete	Promote/ demote	Record performance	Grant/ revoke	Kick/ ban
Common	Session owner	■	■	■	■	■	■	■
Common	Moderator	□	□	■	■	□	■	■
Common	Audience	□	□	□	□	□	□	□
Performance	Performer	■	■	□	□	□	□	□
Performance	Technician	■	■	■	□	■	□	□
Game/Free jam	Contributor/Player	■	■	■	□	□	□	□
Lesson	Student	■	□	□	□	□	□	□

Table 2. Adjustable Role-rights matrix for *WAM Jam Party* session contexts. ■ = right granted; □ = right not granted.

tening modes or mechanisms to temporarily group nodes for private monitoring are also considered to avoid interferences between each user's listening setup. Simplifying the creation of private listening contexts would reduce the need to rewire nodes to different outputs while iterating.

Moreover, prototypes for VRMIs such as the virtual drum kit already implement a form of personal territory (gestures remain local until the user records) [25]. Generalizing this into a *copy/push* model, where any user can audition a personal copy of a subsystem with acoustic privacy before pushing it to the shared session could significantly reduce accidental disruptions during iterative creation.

G6 – Keep ownership and editing policies configurable: Session sharing mechanisms should preserve ownership of saved installations, preventing other users from altering or destroying existing configurations. As in *Virtuoso*'s Community Library system, other users should be allowed to make their own copy of a session. The exact permissions given to a user (defaulting from the permission matrix in section 4.3) could be made configurable by session owners. To avoid excessive complexity, this will not be a priority.

Additionally, node groups could be associated with a simple on / off locking mechanism for users to activate on their owned nodes. Similarly, users performing critical tasks such as adjusting continuously modifiable parameters or recording patterns with VRMIs could temporarily lock said parameters or nodes to prevent conflicting edits.

4.2. Additional design mechanisms

Beyond the guidelines discussed above, several complementary mechanisms could further support collaboration in *WAM Jam Party* and similar CMM SVEs.

Audience participation and lightweight interactions: game-inspired mechanisms may help integrate audiences more actively into collaborative performances. Rather than assigning editing permissions, spectators could interact through lightweight actions that do not interfere with the musical structure (e.g., triggering visual effects, influencing reward systems, or reacting through gestures, competition modes). Such mechanisms could strengthen engagement while preserving the stability of the shared musical environment. Audience-related roles could also benefit from dedicated accessible interactions requiring no permissions to favor inclusion, such as "dancing with others" interactions acting upon the shared musical product.

Cross-platform roles: Because WebXR environments can be accessed from both HMDs and desktop browsers, role-specific interfaces could support different responsibilities. For instance, moderators or technicians might supervise sessions or manage permissions from desktop interfaces, while performers interact from an HMD.

Distributed parameter control: Control over musical parameters could also be distributed across participants. Previous studies suggest that such designs can support collaboration when responsibilities remain clear and users can perceive the effects of their actions.

4.3. Proposed role distribution by use context and permission matrix

Intended as a design proposal, Table 2 displays a theoretical context-sensitive role–rights matrix for *WAM Jam Party*, inspired by multi-role UX perspectives [1]. This matrix serves as a starting point for implementation and will require iterative refinement through user studies to assess its impact on group dynamics. Roles are organized according to typical collaboration contexts (performance, lesson, free jam), and each column indicates whether a given role can perform a specific action on the shared musical structure. The matrix reflects varying levels of authority, from read-only audience roles to full-control session owners. Intermediate roles such as performer, contributor / player, or technician represent different degrees of interaction with the shared patch depending on the session context. Note that in our design, users who can add nodes implicitly retain the right to edit or delete nodes they created.

In practice, moderation privileges do not necessarily correspond to a participant's musical role. A moderator may simultaneously act as a performer or contributor during the session. For this reason, we argue that moderation capabilities should be considered as an additional privilege layer rather than a distinct social role. In this model, a participant may combine a base role (e.g., performer, contributor, student) with moderation privileges such as managing access rights or removing disruptive users. This separation avoids promoting moderators to full session owners while still enabling effective session supervision.

We considered further specializing the distribution with subroles but preferred to keep the distribution simple, hence the common owner role for each highest authority level, as teachers or game masters didn't need further specification regarding permissions. Furthermore, the node-based design of *WAM Jam Party* could have complicated concrete implementations of such a specific subdivision.

Different collaborative contexts motivate different role distributions. Three representative scenarios illustrate how the proposed permission matrix may support typical uses:

Performance preparation: Technicians build and configure the installation while a session owner supervises; once stable, performers play instruments while technicians monitor, handle spatialization and mixing.

Music lesson: A teacher guides students through patch construction; students manipulate their own nodes but cannot alter the global structure, while observers join with read-only access.

Open collaborative jam: Participants freely add and modify nodes while self-organizing. Default symmetric permissions combined with awareness cues may be sufficient without strong authority structures.

5. DISCUSSION

Although this study is based on a limited corpus of VR-first CMM systems, it highlights several recurring patterns regarding roles, territories and sharing. The analysis identifies a set of role-related mechanisms that appear across systems and contexts. These include the emergence of organizational roles around performances, the importance of spatial territories in immersive collaboration and the importance of awareness cues in maintaining group flow.

This territorial dimension is closely related to the awareness and engagement mechanisms described in Section 3.5. Workspace awareness frameworks emphasize that collaboration depends on participants' ability to understand who is acting on what, and where. In immersive environments, this awareness is mediated through spatial proximity, avatar embodiment, and visible interaction cues. Roles, ownership indicators, and activity feedback therefore contribute not only to conflict prevention but also to the construction of mutual engagement and group flow.

While the proposed guidelines and permission matrix discussed in Sections 3.6 and 4.3 aim to reduce friction, their actual effectiveness in preserving creative flow remains to be empirically validated. The mismatch between rigid permissions and the fluid nature of improvisation is a structural tension; therefore, our model should be viewed as a proposition for informed exploration.

More broadly, immersive patch-based systems blur the distinction between instrument and environment. In traditional settings, the instrument is a bounded object with a clear owner. In collaborative VR patching environments, the instrument becomes a shared, evolving structure that multiple users can modify. Roles and permissions therefore operate not only as access control mechanisms but as structuring elements of the instrument itself. From this perspective, designing role systems is inseparable from designing the instrument's interaction model.

Another aspect that may become relevant as collaborative VR music platforms mature is user authentication and data management. If persistent user profiles and shared sessions are introduced, questions related to identity management, privacy and data ownership will emerge. Future iterations of *WAM Jam Party* may therefore need to integrate authentication mechanisms while preserving the openness of collaborative sessions.

Lastly, the availability of an open-source prototype provides an opportunity to investigate these questions empirically. Future work could examine how territorial mechanisms, ownership models, and role configurations affect collaboration dynamics, creative flow, user experience and engagement. In particular, comparing embodied private interaction (e.g., instrument-local control such as the drumkit) with explicit permission-based public coordination could help clarify how spatial interaction and formal role structures complement or substitute each other.

The design guidelines, additional mechanisms and permission matrix proposed in Sections 3.6, 4.2 and 4.3 represent a first operationalization of these principles, but their broader implications remain to be explored empirically.

These observations suggest that roles, territories and embodied interaction form an interconnected design space. Rather than being treated as purely technical coordination mechanisms, role design should be understood as a central aspect of interaction design that shape how collaborative musical experience unfolds in immersive environments.

6. CONCLUSION

This review analyzed recurring role-related patterns across collaborative VR music systems (**RQ1**) and adjacent CMM systems, and distilled a set of design principles supporting collaboration and creative flow while reducing the need for a formal role distribution system (**RQ2**). Through a case study applied to the open-ended CMM SVE *WAM Jam Party* (**RQ3**), these principles are applied and complemented by additional mechanisms extrapolated from the review as well as a preliminary permission model proposition that can be refined in future work.

Beyond its practical implications, this work also highlights the need to further investigate how role models shape musical agency, authorship, and collaborative experience.

The results of this study are currently being implemented in the open-source WebXR prototype of *WAM Jam Party*. This implementation will serve as a testbed to evaluate the proposed role distributions and permission mechanisms in realistic collaborative music scenarios. One planned scenario involves the preparation of a collective musical performance with approximately ten musician–composer participants. This process will include an initial exploration phase, followed by the collaborative construction of a shared musical setup and subsequent rehearsal sessions of the intended composition. These stages will allow us to observe how role configurations support coordination, creative flow, and conflict management during extended collaborative work. Such experiments will provide empirical feedback to refine the proposed guidelines and matrix and assess their applicability to VR CMM systems. This scenario is inspired by a prior collaborative experiment conducted with Erasmus Mundus DIGICREA [10] Master students from University Jean Monnet, France, in December 2025, using the *Sequencer Party* platform. This experiment will serve as a useful reference for analyzing collaborative dynamics in *WAM Jam Party*.

[10] https://master-digicrea.univ-st-etienne.fr/en/index.html

7. ACKNOWLEDGMENTS

This work was supported by the French government through the France 2030 investment program, operated by the French National Research Agency (ANR), under grant ANR-17-EURE-0004 (UCA DS4H).

8. REFERENCES

[1] L. N. Azevedo, D. G. Trevisan, and E. W. Clua. Supporting Designers of Immersive Public Installations: Insights from a Multi-Role UX Perspective. In *Anais do XXIV Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais (IHC 2025)*, pages 1205–1230, Brasil, Sept. 2025. Sociedade Brasileira de Computação - SBC.

[2] T. Blaine and S. S. Fels. Contexts of collaborative musical experiences. In *Proceedings of the 2003 Conference on New Interfaces for Musical Expression (NIME 2003)*, pages 129–134. NIME, June 2003.

[3] A. Boem, D. Dziwis, M. Tomasetti, S. Etezazi, and L. Turchet. "It Takes Two" - Shared and Collaborative Virtual Musical Instruments in the Musical Metaverse. In *2024 IEEE 5th International Symposium on the Internet of Sounds (IS2)*, pages 1–10, Erlangen, Germany, Sept. 2024. IEEE.

[4] A. Boem and L. Turchet. Musical Metaverse Playgrounds: exploring the design of shared virtual sonic experiences on web browsers. In *2023 4th International Symposium on the Internet of Sounds*, pages 1–9, Pisa, Italy, Oct. 2023. IEEE.

[5] B. A. Bradley. The next act: Evaluating tools and envisioning the future of VR theatre. In *Journées d'informatique théâtrale*, Avignon, France, Oct. 2024. hal-05176640.

[6] L. Bruns, B. Saurbier, T. M. Voong, and M. Oehler. Presence and Flow in Virtual and Mixed Realities for Music-Related Educational Settings. In *2024 IEEE 5th International Symposium on the Internet of Sounds (IS2)*, pages 1–7, Erlangen, Germany, Sept. 2024. IEEE.

[7] N. Bryan-Kinns and F. Hamilton. Identifying mutual engagement. *Behaviour & Information Technology*, 31(2):101–125, Feb. 2012.

[8] N. Bryan-Kinns and P. G. T. Healey. Daisyphone: Support for remote music collaboration. In *Proceedings of the 2004 Conference on New Interfaces for Musical Expression (NIME 2004)*, pages 27–30, Hamamatsu, Japan, June 2004.

[9] M. Buffa and T. Burns. Real-Time Collaborative Music Creation on the Web: exploiting Web Audio Modules for Interactive Performance and Composition. In *2025 IEEE Symposium on Computers and Communications (ISCC)*, pages 1–8, Bologna, Italy, July 2025. IEEE.

[10] M. Buffa, T. Burns, C. Ferrari, J. Kleimola, O. Larkin, J. Lebrun, S. Letz, Y. Orlarey, S. Ren, and S. Yi. Ten Years of Web Audio Modules: Audio Plug-ins for the Web. *Journal of the Audio Engineering Society*, 73(12):804–826, Dec. 2025.

[11] M. Buffa, M. Winckler, Q. Escobar, S. Demont, A. Hofr, and A. Mir-Sadjadi. Interactive Audio Sculpting: Plugin Customization and UI Affordances in Immersive Environments. In *2025 IEEE 6th International Symposium on the Internet of Sounds (IS2)*, pages 1–10, L'Aquila, Italy, Oct. 2025. IEEE.

[12] M. Buffa, M. Winckler, Q. Escobar, S. Demont, A. Hofr, and A. Mir-Sadjadi. Using Web Audio Plug-ins in 3D environments. In *Proceedings of the 2025 Web Audio Conference (WAC 2025)*, Paris, France, 2025.

[13] C. Çakmak, A. Çamcı, and A. G. Forbes. Networked virtual environments as collaborative music spaces. In *Proceedings of the 2016 Conference on New Interfaces for Musical Expression (NIME 2016)*, pages 106–111, Brisbane, Australia, July 2016.

[14] A. Çamcı, A. Willette, N. Gargi, E. Kim, J. Xu, and T. Lai. Cross-platform and Cross-reality Design of Immersive Sonic Environments. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 127–130, June 2020.

[15] D. Dziwis, H. Von Coler, and C. Pörschmann. Orchestra: A Toolbox for Live Music Performances in a Web-Based Metaverse. *Journal of the Audio Engineering Society*, 71(11):802–812, Nov. 2023.

[16] C. A. Ellis, S. J. Gibbs, and G. Rein. Groupware: Some issues and experiences. *Communications of the ACM*, 34(1):39–58, 1991.

[17] A. Z. Fanani, K. Hastuti, A. M. Syarif, and A. R. Mulyana. Rule-based Interactive Learning Application Model on How to Play Music Instruments. *International Journal of Emerging Technologies in Learning (iJET)*, 15(15):52, Aug. 2020.

[18] R. Fencott and N. Bryan-Kinns. Hey Man, You're Invading My Personal Space! Privacy And Awareness In Collaborative Music. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME 2010)*, pages 198–204, Sydney, Australia, June 2010.

[19] M. Gioiosa, F. Avanzini, L. A. Ludovico, S. Brambilla, and L. Ripamonti. Sound Blocks VR: An Accessible Virtual Reality Musical Instrument. In H. Plácido Da Silva and P. Cipresso, editors, *Computer-Human Interaction Research and Applications*, volume 2371, pages 316–330. Springer Nature Switzerland, Cham, 2025. Series Title: Communications in Computer and Information Science.

[20] C. Gutwin and S. Greenberg. A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Computer Supported Cooperative Work (CSCW)*, 11(3-4):411–446, Sept. 2002.

[21] M. Kaltenbrunner, S. Jorda, G. Geiger, and M. Alonso. The reacTable*: A Collaborative Musical Instrument. In *15th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE'06)*, pages 406–411, Manchester, UK, 2006. IEEE.

[22] L. Men and N. Bryan-Kinns. LeMo: Supporting Collaborative Music Making in Virtual Reality. In *2018 IEEE 4th VR Workshop on Sonic Interactions for Virtual Environments (SIVE)*, pages 1–6, Reutlingen, Mar. 2018. IEEE.

[23] L. Men, N. Bryan-Kinns, and L. Bryce. Designing spaces to support collaborative creativity in shared virtual environments. *PeerJ Computer Science*, 5:e229, Nov. 2019.

[24] L. Men and D. Zhao. Designing Privacy for Collaborative Music Making in Virtual Reality. In *Audio Mostly 2021*, pages 93–100, virtual/Trento Italy, Sept. 2021. ACM.

[25] A. Mir-Sadjadi, M. Buffa, and M. Winckler. Embodied Virtual Instruments in Web-Based Multi-User VR: A Case Study with a 3D Drum Kit and Web Audio Modules. In *Proceedings of the 9th Web Audio Conference (WAC)*, Nov. 2025.

[26] PatchXR. Patchworld (version 77.1.6) [software]. `https://www.patchxr.com/`, 2022. [Online]. Available: https://www.patchxr.com/.

[27] Really Interactive AB. Virtuoso (version 1.1.8) [software]. `https://fasttravelgames.com/games/virtuoso`, 2022. [Online]. Available: https://fasttravelgames.com/games/virtuoso.

[28] A. Syukur, P. N. Andono, and A. M. Syarif. Gamelan Land: A Multiplayer Virtual Reality Game based on a Social Presence Approach. *Journal of Metaverse*, 4(1):1–10, June 2024.