

Mallets in Reality: Avatar-Guided Learning on a Mixed Reality Marimba Digital Twin

TeAiris Majors
Creative Music Technology Lab
Georgia Institute of Technology
Atlanta, USA
teairis@gatech.edu

Alexandria Smith
Creative Music Technology Lab
Georgia Institute of Technology
Atlanta, USA
alexandria.smith@gatech.edu

Abstract—Access to full-sized mallet percussion instruments is often constrained by cost, space, and availability, limiting opportunities for practice. This paper presents *Mallets in Reality*, a mixed reality marimba environment pairing a one-to-one digital twin of a five-octave marimba with avatar-guided instruction and color-based visual supports. In a three-session exploratory study, eight university students, most with limited or no mallet percussion experience, learned through avatar-led call-and-response, falling-note visualization, and hybrid mixed reality and monitor instruction. Participants adopted foundational techniques, including wrist motion and hand orientation, and played physical scales accurately ($M = 0.967$ to 0.992), while the mixed reality song activity was lower and more variable ($M = 0.767$). Scale visualization tools were valued most consistently, and tactile feedback, mallet rebound, and force control most distinguished the virtual instrument from the physical one. With no pre-test or control condition, the study does not quantify learning gains, but it identifies where virtual practice complements rather than replaces the physical instrument.

Index Terms—Mixed Reality, Music Education, Mallet Percussion Pedagogy, Digital Twin, Embodied Learning

I. INTRODUCTION

As a band director at a Title I¹ high school in the U.S., the first author encountered firsthand the limitations of access to mallet percussion instruments. A limited budget prompted building a vibraphone and marimba from scratch to support student practice and development. This experience revealed a persistent challenge: access to full-sized mallet percussion instruments is often constrained by cost, space, and availability [2], [3]. Portable alternatives such as bell kits provide only partial access, lacking the range and spatial complexity needed to develop technique, coordination, and musical fluency. As a result, many percussionists have limited opportunities for sustained, embodied practice outside of rehearsal contexts, which remains central to development well beyond the beginner stage [4].

At the same time, the emergence of extended reality (XR) and the Musical Metaverse introduces new possibilities for music performance and pedagogy [5], [6]. Embodied presence in these environments may strengthen motivation and

engagement [7], [8], while also enabling more accessible and inclusive forms of musical participation [2], [9], and such systems point toward future possibilities in networked collaboration [10]. For instrumental learning, these systems generally take one of two forms: fully virtual instruments rendered in a head-mounted display and played with tracked hands or controllers, or physical instruments augmented with virtual guidance, feedback, or instruction [11], [12].

Mallet percussion presents a particularly demanding case for fully virtual instruments. It requires full-body coordination, spatial navigation across a multi-octave layout, and the development of motor patterns tied directly to physical geometry [13]. Supporting this interaction in virtual spaces demands not only accurate sound reproduction, but also spatial fidelity, interaction realism, and purposeful feedback [11]. However, while educational music XR systems continue to advance technically, instruction within them is dominated by cues that indicate which note to play, with far fewer systems emphasizing technique through full-body avatar imitation of a performer [12].

In educational XR environments, before addressing challenges such as networking and multi-user interaction, it is necessary to first understand how learners engage with virtual instruments and whether those experiences carry beyond the headset. *Mallets in Reality* is therefore designed as a single-user building block toward such environments rather than a networked Musical Metaverse system in itself [6]. This paper argues that instruments within XR environments, designed as digital twins paired with pedagogical framing, can expand access to musical training while preserving the physical and technical demands that performance on the real instrument requires.

This work examines that argument in practice through the creation of *Mallets in Reality* and an exploratory study aimed at further developing learning environments for music, particularly in embodied interaction, digital twin design, and mallet percussion pedagogy. It also positions instructional design as a promising foundation for future Musical Metaverse learning experiences, with avatars offering a form of guidance closer to in-person instruction than flat-screen media allows. Three questions guide this investigation: **RQ1**: How do learners with little or no mallet percussion experience engage with a

¹Title I, Part A of the Elementary and Secondary Education Act provides supplemental financial assistance to school districts for children from low-income families in the U.S., with the aim of closing educational achievement gaps [1].

virtual instrument when practice is self-directed? **RQ2:** Do the techniques and patterns introduced in mixed reality appear when learners return to the physical instrument? **RQ3:** Which forms of instructional support help learners most, and where does the virtual experience fall short of the physical one?

The paper proceeds by situating *Mallets in Reality* within prior work on percussion instruments in XR, embodied musical learning, and the Musical Metaverse. It then details the design of the system, from the digital twin and its interaction model to the avatar instruction and visual supports built around it. Finally, it frames and outlines the three-session study, reports its findings, and discusses what they suggest about where virtual practice supports learning and where the physical instrument remains necessary.

II. RELATED WORK

Research on extended reality for music has explored virtual instruments, embodied approaches to musical learning, and shared performance spaces where musicians interact through avatars. Each of these areas informs a different part of this work: what a virtual mallet instrument must support, how instruction within it can be delivered, and what larger environments it builds toward.

A. Percussion Instruments in XR

Existing applications of extended reality for percussion provide important foundations for virtual instrument interaction but remain limited in their support for mallet percussion. Academic work in this space has concentrated largely on drum set and single-surface instruments: Pinkl et al. developed a multimodal drumming education tool in mixed reality [14], and Banquero et al. evaluated mixed reality drum practice using a commercial drum controller alongside a head-mounted display [15]. Commercial applications show a similar emphasis: *Percussive VR* supports virtual mallet instruments alongside static notation displays [16], and *Paradiddle* centers on rhythm-based drum set practice, with bars for virtual mallet instrument playback [17]. Although these systems demonstrate how XR can support engagement in percussive contexts, they do not explicitly influence technique or include guided performance methods for mallet percussion. The demands differ: where drum set practice is largely stationary, mallet performance depends on navigating a multi-octave layout, and striking accuracy is closely tied to the control of limb movement [13]. This pedagogical gap motivates the development of *Mallets in Reality*.

B. Embodied Interaction and Musical Learning in XR

Embodied interaction plays a key role in musical learning in extended reality environments, where performance can emerge through coordinated physical movement, spatial awareness, and real-time feedback [18]. Yu et al. examined how embodiment and visual cues together affect instrument learning in virtual reality [19]. Research in virtual percussion has examined how auditory and haptic feedback shape the perception of a virtual drum surface [20] and how learners progress in rhythm

performance through VR drum practice [21]. Pinkl and Cohen explored virtual co-embodiment in VR drumming pedagogy, where learners shared avatar control with an autonomous exemplar [22]. Imitation-based approaches like this remain uncommon: a recent review of 74 XR musical instrument systems identified only five that taught through imitation of an expert, most by superimposing an expert's hands onto the instrument rather than presenting a full-body avatar [12]. Hackl and Anthes [23], Guo et al. [24], and Johnson et al. [25] further explored overlaying spatial and visual feedback onto physical instruments to support teaching and learning, demonstrating the pedagogical value of blending physical and virtual layers. Fernandez et al. took a related approach, using an AR virtual character to deliver real-time performance feedback to piano learners [26].

C. Musical Metaverse and Virtual Avatars

The Musical Metaverse extends XR-based music systems into shared, networked environments where performance and interaction unfold across physical and virtual spaces. Turchet et al. describe this development through the convergence of extended reality and the Internet of Musical Things, emphasizing distributed presence, low-latency communication, and persistent virtual environments as core system characteristics [6], [10], [27]. Within these environments, avatars often serve as a primary medium of interaction, enabling users to engage with instruments and other performers through embodied representation. Hopkins et al. demonstrate how remote performers can drum together through avatars rendered in augmented reality [28], and have since extended this to AI-driven virtual musicians performing alongside human players [29], highlighting new forms of musical participation within shared XR spaces.

Comparatively less attention has been given to how these systems support musicians as supplementary resources. Understanding the importance of representational figures, how users engage with virtual instruments, and whether these interactions translate to physical performance remains underexplored. Because technique is also learned through movement across the instrument itself, whether virtual practice carries over to the physical instrument becomes a central question.

III. SYSTEM DESIGN AND IMPLEMENTATION

Mallets in Reality is designed not only as a virtual instrument, but as an instructional system grounded in embodied learning and representation. This section begins with the instructional framing behind the system and the role of avatars within it. It then describes the instrument itself, including its physical dimensions, bar and mallet interaction, feedback, and sound production. Finally, it presents the visual tools and avatar instruction built around the instrument, along with options for learners to adapt the system to their own needs.

A. Instructional Framing and Representation

Mallet percussion instruction is highly visual and embodied, relying on nuanced demonstrations of posture, grip, stroke

type, and spatial navigation across the instrument. Embodied avatars in virtual environments allow this instruction to be recorded, replayed, and spatially aligned with the learner’s own instrument configuration, affording demonstrations that a learner can revisit as often as needed.

Equally important is the role of avatars in expanding representation within the field. Public education in the United States remains demographically imbalanced. National data from a 2020–2021 study indicate that roughly 80% of educators identify as White, while student populations have become significantly more diverse, highlighting a persistent representation gap [30]. Within this context, mallet percussion instruction often reflects a similarly narrow set of visible models in both educational materials and performance spaces [31]. This system intentionally positions avatars as a flexible representational layer, with the instructing figure drawn from an interchangeable avatar library, allowing for the inclusion of diverse identities and instructional voices. In doing so, this work responds to broader issues of belonging by enabling learners to see themselves reflected within instructional content. The study reported later in this paper used a single avatar and voice throughout for consistency, so this representational capacity is expressed as a design affordance rather than something the study evaluated.

B. Physical Dimensions and Spatial Fidelity

In the first phase of development, the marimba was selected as the primary instrument to establish a framework for realistic interaction and visualization that can extend to other mallet instruments. A one-to-one dimensional mapping between the physical and virtual instrument (Fig. 1) preserves spatial continuity and supports accurate motor transfer between environments. The system is modeled as a digital twin of a full-sized five-octave marimba informed by a measurement specification reference [32].

The virtual bars were modeled in Blender, beginning from the dimensions of a C2 bar and scaled across the five-octave range to approach a reference C7 bar. Bars were arranged in the chromatic pattern used for keyboard instruments so that spatial awareness developed on the virtual instrument would transfer naturally to common physical instruments. The completed model was imported into Unity, where each bar was procedurally renamed by MIDI note number (36–96, C2–C7), allowing any bar to be referenced by name throughout the system.

While resonators are acoustically significant in physical instruments, they are omitted² in the virtual environment.

C. Bar Interaction and Articulation

To approximate acoustic responsiveness, three articulations are defined: normal, dampened, and node, corresponding to open resonance, immediate damping, and nodal striking techniques [33]. All three are implemented using callback methods of Unity’s collider class.

²Resonators are excluded to improve visibility in pass-through mode and support safe physical navigation while preserving spatial relationships.



Fig. 1. This figure showcases the one-to-one dimensional fidelity between the virtual and physical marimba instruments. The virtual instrument is positioned directly above the physical instrument.

A normal strike is triggered on `OnTriggerEnter`. The dampened articulation calls a coroutine from `OnTriggerStay`, which waits briefly before issuing a Note Off event, replicating a bar struck and immediately stopped. Because a normal stroke leaves the collider quickly, `OnTriggerExit` cancels this coroutine, so damping occurs only when the mallet remains in contact with the bar. The node articulation uses an additional collider positioned slightly above the node, which activates an audio high-pass filter on entry to resemble the attenuated harmonics produced when a bar is struck at its node.

To address limitations in Unity’s predictive collision handling, particularly during fast rhythmic passages, each bar’s collider is narrowed and centered. This minimizes unintended double-triggering and false collisions from neighboring bars during rapid lateral movement, while the width of the mallet head keeps ordinary striking reliable. This adjustment was calibrated for interaction response and does not model the relationship between strike position and tone quality.

D. Mallet Interaction

Unlike comparable percussive XR systems, which attach mallets to the user’s hands, mallets in this environment are not pre-positioned. Users freely grab mallets, after which the virtual hands conform to the recommended grip technique. This approach supports guided skill development while maintaining user agency, allowing novice and experienced players alike to model proper technique. Future iterations will allow this constraint to be relaxed for alternative playing styles.

E. Haptic and Visual Feedback

Haptic feedback is implemented using the controller’s internal vibration method, with amplitude mapped to the scaled velocity magnitude obtained at the moment a bar is struck. Unlike audio-driven haptics authored in Meta Haptics Studio, this approach directly reflects collision intensity, providing responsive feedback that reinforces dynamic control [34]. To further support dynamic awareness, each mallet includes a visual meter indicating strike intensity (Fig. 2).

The calibrated dynamic range encourages appropriate playing force, promoting technique development while discourag-

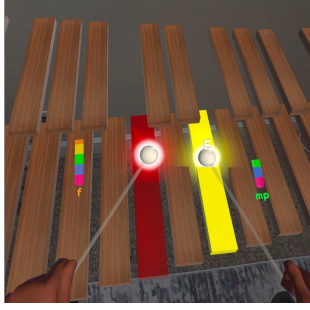


Fig. 2. This figure demonstrates the dynamics visual meter attached to each virtual mallet.

ing excessive impact that could transfer poorly to physical instruments.

F. Sound Production

An initial sound engine used sixty-one chromatic marimba samples mapped by MIDI note number and triggered through Unity’s collider callbacks. Latency was acceptable even during rapid repeated strokes, but the workflow scaled poorly: each additional instrument required its own time-consuming sampling, formatting, and in-editor referencing, and build size grew with every instrument added.

The current system replaces this with a SoundFont-based approach integrated through a FluidSynth backend [35] and accessed using the Maestro MIDI Player Toolkit [36]. SoundFont files use keymapping to spread a limited number of recorded samples across multiple notes, conserving memory while maintaining a natural sound, and allow multiple instrument presets to be compiled into a single soundbank. This system uses the MalletsSectionGM SoundFont [37], a 56 MB bank containing celesta, glockenspiel, vibraphone, marimba, and xylophone patches. A `MidiStreamPlayer` component interprets each strike as an `MPTKEvent`, issuing a `Note On` command with the note number taken from the struck bar’s name and velocity derived from the observed impact. Running FluidSynth on a separate processing thread supports stable low-latency performance, a critical requirement for interactive musical systems [10].

G. Visuals and Assistive Tools

Beyond the instrument itself, the virtual environment affords visual supports that a physical instrument cannot provide. By default the instrument is set to marimba mode, displaying a full five-octave range with a rosewood texture. Upon striking a bar with the virtual mallets, the color of the rosewood texture is temporarily changed and follows the Boomwhackers chromatic color scheme [38], a well-established visual signifier for pitched percussion in educational settings; color-based notation has been used elsewhere to reduce the symbolic demands of early instruction [39]. Pitch labels are positioned above the bars near the far end, outside the normal striking zone, and are revealed after a bar is struck, providing reminders of the notes currently in focus while encouraging memorization of the pitch layout rather than continuous reliance on labels.

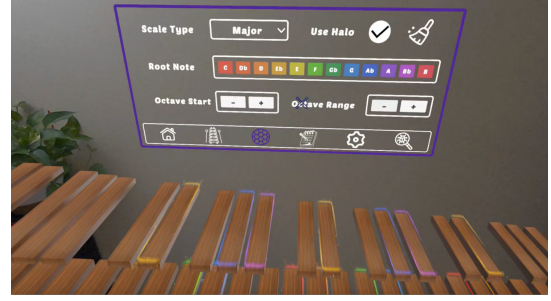


Fig. 3. This figure showcases the assistive scale tool user interface and halo visualization option utilized in Session 3.

Localized text reminders appear near the player’s wrist when mallet orientation deviates from the recommended position, prompting learners on mallet orientation and playing position.

The ability to change bar color provides a guide system for musical concepts such as scales, chords, and melodic patterns. Within the tools menu, users select a scale type and set its tonal center by choosing a root note, along with a starting octave and octave range. Two visualization modes are provided (Fig. 3):

Normal: All bars within a selectable range are illuminated to indicate scale notes. Striking a highlighted note returns the bar to its unlit state.

Halo: All bars within a selectable range are outlined with an illuminated border created with a Unity particle component to indicate scale notes. The halo remains visible after a note is struck.

MIDI files can also be played back and visualized on the instrument, presenting pitch and rhythm as vertically moving cues that descend toward the corresponding bar, color-coded consistently with the virtual marimba.

H. Avatar Instructional System

The instructional system consists of a full-body avatar animation recorder, object transform recorder, and voice recorder unified into a single component, enabling performance gestures and verbal guidance to be captured as cohesive, reusable assets. An Animation Player component sequences these recordings using state-based logic and dynamically adapts them to each user’s instrument height and scale. This adaptation ensures that the avatar remains able to interact with the virtual instrument during demonstrations regardless of the learner’s configuration. All instrument interactions described above hold true when performed by the avatar, so its strikes produce the same articulation, sound, and visual responses as the learner’s own.

The instruction flow is purposefully choreographed so that visual cues align with the instructional content they support. When the avatar strikes a bar, the learner’s controller corresponding to the hand in use vibrates, reinforcing sticking guidance during demonstration. On launch, a virtual avatar briefly introduces the interface, after which learners may engage in a guided lesson covering foundational mallet percussion concepts or begin playing the instrument immediately.

I. Flexible Design and User Autonomy

The system supports both seated and standing use, thumb-stick locomotion for moving between ranges of the instrument, and global instrument scaling to accommodate different user needs and physical contexts. Scaling the full bar layout accommodates differences in reach, resembling the options available on student model instruments. Although most mallet percussion instruments include some form of height adjustment, it often goes unchanged, whether because learners are unaware it is an option, lack the time, or are not permitted to adjust it for safety reasons. In this environment, instrument height is automatically set from an estimated hip level derived from controller position and can be adjusted at any time with a button, with all interface elements updating accordingly. These features support autonomous learning by enabling users to explore the instrument at their own pace and within their own physical constraints.

IV. STUDY DESIGN

A user study, approved by the Institutional Review Board at the Georgia Institute of Technology, was conducted with informed consent from all participants. Across three instructional sessions, the study examined how learners engaged with *Mallets in Reality* and how musical and motor skills introduced there carried into performance on a physical marimba [18]. Instruction was delivered through several approaches across the sessions: avatar-led call-and-response, researcher-led demonstration at the physical instrument, a falling-note visualization activity, and hybrid instruction pairing the mixed reality environment with a monitor display. These were examined as exploratory alternatives rather than compared under matched conditions. Particular attention was given to technique adoption, scale performance, learner engagement, and participants' experiences moving between virtual and physical contexts, corresponding to the research questions outlined in Section I.

A. Population

Ten university students participated in the study and represented a range of academic backgrounds across the arts, sciences, and engineering disciplines. Participants were recruited through university communication channels and completed a pre-study questionnaire before participation. Two participants reported prior mallet percussion experience, while the remaining participants described limited or no experience with mallet percussion, with general musical backgrounds ranging from limited formal training to more than a decade of instrumental practice (Table I). Although all participants reported some prior VR exposure, several described only brief encounters, such as a short demonstration or trial session. Eight participants completed all three study sessions and were included in the final analysis; the two who did not complete the full protocol were excluded from the reported results.

B. Procedure

The study took place in a controlled setting equipped with a physical five-octave marimba and a Meta Quest 3 headset

TABLE I
PARTICIPANT CHARACTERISTICS ($N = 8$)

Characteristic	Participants
Gender	5 male, 3 female
Age (years)	19–31 ($M = 24.1$, $SD = 3.9$)
Reads Western notation	8 / 8
Prior mallet percussion experience	2 / 8
Prior VR/XR experience	Limited; mostly brief or occasional

running all session builds of *Mallets in Reality*, with the Quest 3 controllers serving as virtual mallets during all mixed reality interactions. For all sessions, the global scale of the virtual marimba was fixed to match the dimensions of the physical marimba used in the study, preserving the one-to-one correspondence described in Section III-B. An aftermarket head strap was used to improve comfort and accommodate participants wearing corrective lenses. All sessions were recorded through the headset and with an external camera. Before each session, participants received an overview of the study procedures, the mixed reality environment, and the physical instrument.

Across three sessions, participants progressed through cycles of instruction, self-directed practice, independent exploration, and performance assessment. Instruction in mixed reality was delivered primarily by the avatar, while instruction on the physical instrument was given live by the researcher. During assessments, the researcher specified only the material to be performed; participants completed each assessment independently, without coaching or correction. No standard notation was used at any point in the study; instruction relied on rote demonstration and the visual supports described in Section III-G.

C. Session 1

The first session introduced participants to the study environment and foundational mallet percussion concepts. The virtual avatar led an overview of common mallet instruments, their sounds, and the keyboard layout of the marimba in relation to the piano, supported by in-environment video and animation, as shown in Fig. 4. An introduction to virtual controls, mallet interaction, and instrument height adjustment followed, with the wrist-orientation reminders described in Section III-G reinforcing hand orientation and playing position.

The avatar then led a call-and-response sequence covering striking mechanics, alternating strokes, and the C-major scale. Rote instruction was selected to support accessibility across varying levels of musical experience [40]. Participants were given up to fifteen minutes to freely explore the virtual instrument, after which they transitioned to the physical marimba and performed the C-major scale.

D. Session 2

The second session focused on rhythmic performance, chromatic navigation, and movement across the instrument. Participants began with an interactive song activity using the falling-note visualization described in Section III-G, with a MIDI



Fig. 4. This figure showcases the introduction to mallet instruments demonstrated in Session 1 to relate the marimba key layout to a piano keyboard layout.

score authored by the first author from melodic themes of “Anxiety” by Doechi and “Somebody That I Used to Know” by Gotye featuring Kimbra, followed by a brief stretching routine presented as a pre-recorded video in mixed reality to encourage healthy playing habits [41].

Participants then transitioned to the physical marimba for researcher-led instruction on the chromatic scale across the full five-octave instrument, including note grouping, hand shifting between manuals, and foot positioning. They returned to the mixed reality environment for five minutes of self-directed practice, and the session concluded with a physical assessment of chromatic-scale performance.

E. Session 3

The third session integrated virtual and physical learning. Participants performed on the physical marimba while receiving instruction through the mixed reality system, with the headset feed displayed on a monitor in front of the instrument and audio routed through room speakers. Instruction covered the B-flat and E-flat major scales, after which participants were introduced to the assistive scale tools described in Section III-G and given five minutes of self-directed practice using the scale interface. The session concluded with a physical assessment of the two scales.

F. Measures and Data Collection

Data collection included performance assessments, technique observations, exploration time, and post-study survey responses, supplemented by researcher observations of visual aid use, chromatic navigation, hand-shifting strategies, and movement patterns carried between virtual and physical contexts. Free exploration in Session 1 was timed from its start until participants elected to stop or were asked to stop at the fifteen-minute maximum.

1) *Performance Accuracy*: All assessments were scored by the first author from video recordings: physical assessments from the headset and external camera recordings, and the Session 2 song activity from the headset recording. Major scales were performed one octave ascending and descending without repeating the top note (15 notes each for C major, B-flat major, and E-flat major), and the chromatic scale spanned the full five-octave range ascending and descending (121 notes).

Participants performed each physical assessment at a self-selected tempo, while the song activity (80 notes) was paced by the falling-note visualization. Accuracy was calculated as the proportion of correctly struck notes, and the number of manual-shift errors was also recorded during the chromatic assessment.

2) *Technique Adoption*: Three technique measures (hand orientation, wrist motion, and alternating strokes) were observed during the Session 1 C-major assessment and coded as correct or incorrect for each participant. No case arose in which one hand met a criterion while the other did not.

3) *Post-Study Survey*: Following completion of all study sessions, participants completed a written post-study survey examining learning experiences, instructional effectiveness, engagement, accessibility, headset comfort, and perceived realism, combining rated items with open-ended questions.

G. Analysis

Performance and interaction data were analyzed descriptively. Accuracy is reported as the mean, standard deviation, and range for each task; technique adoption as the proportion of participants demonstrating each measure; and exploration time as the mean, standard deviation, and range. Rated survey items are reported as counts of positive, neutral, and negative responses, since the items used labeled scales of differing lengths. Given the sample size and single-condition design, no inferential statistical tests were conducted.

Open-ended survey responses were used narratively. Responses were read in full, and passages were selected to illustrate participants’ experiences in relation to each research question. No interviews were conducted; all quotations are drawn from written survey responses and are attributed by participant number. No formal coding scheme was applied.

V. RESULTS

A. Technique Adoption and Early Skill Development

Most participants successfully adopted the foundational performance techniques introduced during instruction, as coded from video of the Session 1 C-major assessment (Section IV-F). All eight participants demonstrated correct wrist motion, and seven of eight demonstrated correct hand orientation and alternating strokes. Six of eight participants met all three criteria, and the two exceptions each missed a single measure.

Written reflections referenced the same elements. When asked what they learned during Session 1, participants referenced proper mallet grip, wrist rotation, hand position, and striking technique. One participant reported learning “the correct way to hold the mallets” (P2, Q3.6), while another highlighted “rotating the wrists” as a meaningful takeaway (P3, Q3.6). A third participant specifically noted that the visual reminder system reinforced correct technique, stating “Just reminding me to play palms down!” (P5, Q3.6).

TABLE II
PERFORMANCE ACROSS ASSESSED TASKS ($N = 8$)

Task	Context	Mean	SD	Range
C-major scale	Physical	0.992	0.024	0.933–1.000
Chromatic scale	Physical	0.981	0.020	0.950–1.000
B-flat major	Physical	0.967	0.050	0.867–1.000
E-flat major	Physical	0.975	0.050	0.867–1.000
Song activity	Mixed reality	0.767	0.118	0.475–0.888

B. Performance Across Tasks

Following mixed reality instruction and practice, participants completed performance assessments on the physical marimba, along with a song activity in mixed reality during Session 2. Table II summarizes accuracy for each task as the proportion of correctly struck notes, which remained high across the physical tasks and was substantially lower and more variable for the song activity.

The contrast between these tasks is informative. The physical scale assessments were performed at a self-selected tempo and measured note selection and navigation, while the song activity was externally paced by the falling-note visualization and combined pitch, rhythm, and movement demands. One participant fell below 0.50 on the song activity. High accuracy on the scale assessments should therefore not be read as general competence on the instrument. Manual-shift errors during the chromatic assessment averaged 2.0 per participant (range 0–6), indicating that navigation across the full five-octave range remained demanding even where note accuracy was high. Because no pre-assessment was administered, these values establish a level of performance rather than a measured improvement.

Some participant reflections addressed the transition between virtual and physical performance contexts. One participant noted that after removing the headset and playing a physical marimba, “it felt like I understood the instrument and the only calibration I really had to do was the force I applied to the wooden keys as opposed to the VR keys” (P6, Q3.6). The same participant later reported that “practicing in VR definitely improved my real-world playing” (P6, Q3.11).

Participants also highlighted skills related to movement and navigation across the instrument, with one describing learning “the action of shifting hands when playing chromatic scale; the stepping needed to reach further notes” (P2, Q3.11).

C. Survey Ratings

Following the final session, participants rated 23 items covering learning experience, instructional effectiveness, engagement, accessibility, comfort, and perceived realism, with responses grouped by valence (Section IV-G). Thirteen items drew uniformly positive responses from all eight participants, including the assistive scale interface, perceived realism, and willingness to recommend the system. The remaining ten items, shown in Table III, included at least one neutral or negative response.

Two items warrant attention beyond their distribution. Ratings of the avatar demonstrations spanned the full range of

TABLE III
RATED SURVEY ITEMS WITH NEUTRAL OR NEGATIVE RESPONSES

Item	Pos.	Neu.	Neg.
Call-and-response instruction (S1)	7	1	0
Color-coded notes aided understanding	7	1	0
Falling notes: rhythm and timing	7	0	1
Intuitiveness of mallet interaction	7	0	1
Color aided music theory understanding	6	2	0
Falling notes: pitch identification	6	2	0
Avatar demonstrations of technique	6	1	1
Demonstrations corrected technique	6	1	1
Mixed reality and monitor instruction (S3)	6	0	2
Headset comfort (6 responding)	4	0	2

the scale, from very ineffective to extremely effective, making this the most divided item in the survey. Ratings of the falling-note visualization for rhythm and timing included the single lowest rating recorded anywhere in the survey, consistent with the reduced accuracy observed on that task. Headset comfort was likewise divided, with two of the six participants who answered reporting discomfort.

D. Visual Supports and Learning Aids

Participants frequently referenced visual learning aids when describing their experiences with the mixed reality environment. Commonly cited features included color-coded bars, note labels, falling-note visualizations, scale highlighting, and halo indicators.

Several participants described the color-coding and highlighting systems as particularly helpful for understanding note relationships and scale patterns. One participant identified “the halos and color-coding of some of the mallet keys to show scales” because “it allowed me to follow the visuals and play right away” (P3, Q4.5). Another participant cited “AR features such as alphabetic labels placed on keys, highlighting of keys, [and the] virtual instructor” as major strengths of the application (P2, Q7.3). Other participants further referenced the scale visualization tools introduced during Session 3, noting that repeated practice with color-coded scale notes supported learning and recall of major scale patterns. For example, one participant stated that “seeing the colors greatly helped in playing on the real marimba” (P6, Q3.12).

Responses to the falling-note visualizations introduced during Session 2 were more divided. One participant reported that “falling notes felt like playing a rhythm game, and it was exciting” (P4, Q3.11) and identified “Session 2 falling notes” as one of the most engaging aspects of the experience (P4, Q4.5). Others described difficulty interpreting the cues. The same participant found it “challenging to identify what the ebony note visualization refers to” (P4, Q2.5), while another reported that “indicating the correct bar or pitch to strike could be more evident,” noting that “during the 2nd session, there are more things that you are doing so it is hard to keep track of those cues” (P10, Q3.11). A third described difficulty determining “exactly which part of the rhythm to play correctly based on the visual guide that accompanied the music” (P7, Q2.5). Taken together with the accuracy reported

in Table II, these responses suggest that the visualization conveyed pitch adequately but did not communicate rhythmic placement clearly enough under the combined demands of the task.

E. Engagement and Learner Agency

Following avatar-led instruction in Session 1, participants were given the opportunity to freely explore the mixed reality instrument. With the same instruction and environment available to everyone, exploration times varied considerably, with a mean of 576.4 seconds ($SD = 304.8$), ranging from 150 seconds to the maximum allotted duration of 900 seconds. Three participants continued interacting with the instrument until at or near the time limit.

Several participant reflections described the experience as enjoyable, engaging, and motivating. One participant remarked, “This was a really creative experience that I found myself enjoying much more than I had initially anticipated!!!” (P5, Q7.5), and another, highlighting autonomous exploration, identified “the haptic feedback, color coding of notes, ability to experiment and play freely” as particularly engaging (P6, Q4.5). Some participants additionally reported increased interest in learning mallet percussion, with one noting that the experience “made me very curious to take on the marimba as an instrument that I could actually learn” (P10, Q7.5).

F. Accessibility and Perceived Realism

Several participants discussed accessibility when describing the value of the mixed reality environment, referencing the cost and availability of mallet percussion instruments. One participant identified “Ease of use. Real physical instruments are expensive and hard to gain access to” as a major strength of the application (P5, Q7.3).

Participants also identified areas for improvement related to physical interaction, including mallet weight, impact sensation, and tactile feedback. One participant suggested “maybe more weight to the mallets?” (P6, Q6.6). Another identified the tactile gap most directly, writing that “the vibration in the controller when hitting a note is nice, but the action of bouncing the mallet off of the marimba key to let the note ring is a very specific skill that made the gap between the physical and virtual instrument most apparent” (P3, Q6.5). Participants also referenced accidental strikes, collision sensitivity, and note-registration issues, such as occasionally striking unintended adjacent bars (P3, Q2.6) and visual mallet contact that did not always align with note activation (P5, Q4.6). Despite these observations, participants often characterized the virtual instrument as realistic in relation to its educational purpose, with one describing the experience as “almost as effective as practicing in reality” for developing the movements associated with mallet percussion performance (P2, Q6.5).

VI. DISCUSSION

A. Learning Across Virtual and Physical Contexts

Participants appeared able to move between virtual and physical performance contexts with relatively little reported

TABLE IV
OPEN-RESPONSE SURVEY PROMPTS CITED IN SECTION V

ID	Prompt
Q2.5	Were there any parts of the interface or features that were confusing or difficult to use?
Q2.6	What challenges, if any, did you face while interacting with the virtual tools?
Q3.6	What did you learn during session 1 that you feel will significantly impact your practice sessions?
Q3.11	What did you learn during session 2 that you feel will significantly impact your practice sessions?
Q3.12	What did you learn during session 3 that you feel will significantly impact your practice sessions?
Q4.5	What aspects of the VR experience were most engaging?
Q4.6	What, if anything, detracted from your sense of immersion or engagement?
Q6.5	Do you feel this VR application could replace physical mallet percussion instruments for practice? Why or why not?
Q6.6	What improvements would make the application feel more realistic or effective?
Q7.3	What were the most significant strengths of the VR application?
Q7.5	Is there anything else you would like to share about your experience with the VR application?

difficulty. While the study design does not permit claims of learning gains or causal transfer, the findings are consistent with prior research suggesting that embodied interaction can support learning in XR environments by linking movement, perception, and feedback into a unified activity [18]. In the present study, participants practiced within a one-to-one digital twin of a full-sized marimba before performing on its physical counterpart, keeping spatial relationships and movement patterns largely consistent.

Importantly, the findings suggest that some aspects of performance translated more readily than others. Participant reflections and researcher observations indicated that note locations, scale patterns, hand-shifting strategies, and movement across the instrument were carried between virtual and physical performance contexts. This distinction is noteworthy because mallet percussion performance depends not only on note knowledge but also on coordinated movement across a large spatial layout [13].

At the same time, participants consistently identified elements of physical performance that were not fully reproduced in the mixed reality environment, with one noting that the primary adjustment on the physical instrument was the force applied to the bars and others highlighting the feel of mallet rebound. These observations align with prior work emphasizing the importance of sensory feedback in virtual musical instrument interaction [11], [20].

B. Instructional Supports and Interface Design

Instructional supports were not received equally. The scale visualization tools were among the most consistently praised features of the system. Participants referenced color-coding, highlighting, and halo visualizations when describing how they learned note patterns and navigated the instrument. These

tools appear to have reduced some of the cognitive demands associated with locating notes and recalling scale structures, allowing learners to focus more directly on movement and performance, consistent with prior work on visual augmentation for instrumental learning [23]–[26].

In contrast, responses to the avatar demonstrations were more varied, suggesting that observation of a full-body virtual performer was beneficial for some learners but less meaningful for others. Because participants were not asked why they rated the system differently, the present study cannot determine the source of these differences. Nevertheless, the variation highlights that instructional supports should not be assumed to benefit all learners equally, even when they are grounded in established observational and imitation-based learning approaches [12], [22].

The falling-note visualization presented the greatest challenge. Unlike the scale visualizations, which primarily supported spatial navigation and note identification, the falling-note interface required learners to coordinate pitch recognition, timing, movement, and performance simultaneously. These findings suggest that the challenge of the activity may have stemmed less from the musical material itself and more from the design of the instructional interface.

C. Accessibility and the Role of Physical Instruments

Several participants described the system as a potentially valuable way to engage with mallet percussion when access to physical instruments is limited, consistent with broader discussions of accessibility in music education and XR systems [2], [9]. These findings should be interpreted cautiously, however. The study examined perceptions of accessibility rather than measuring changes in access, and it does not demonstrate that the system increases participation or reduces educational inequities.

Accessibility also extends beyond instruments. While mixed reality may reduce the need to obtain and store large percussion instruments, it introduces new requirements in hardware, physical space, and technical familiarity. Technologies intended to improve access must therefore consider exclusion through unequal access to XR devices, and future work should weigh instrument accessibility alongside the broader digital divide.

At the same time, participants generally described the system as a supplement rather than a replacement for physical instruments, suggesting that mixed reality is most valuable as a complementary resource for early familiarization, self-directed practice, and situations in which physical instruments are unavailable.

VII. CONCLUSION

This study examined how learners with limited mallet percussion experience engaged with a mixed reality marimba environment and subsequently performed on a physical instrument. Across three instructional sessions, participants adopted foundational performance techniques, completed physical performance assessments with high accuracy, and reported mostly

positive experiences with visual learning aids, guided instruction, and opportunities for independent exploration. The findings suggest that mixed reality environments can support early familiarization with instrument layouts, performance movements, and foundational playing techniques. They also highlight the value of instructional design within XR music-learning systems, particularly visual supports that assist navigation and concept development. At the same time, participants identified tactile feedback, mallet rebound, and force control as aspects of physical performance that remained difficult to reproduce virtually.

A. Limitations

These conclusions should be read in light of the study's exploratory design. It included no pre-test, post-test, control group, or comparison condition. Participant improvement therefore cannot be quantified, and the contribution of mixed reality instruction cannot be separated from prior experience, practice effects, or repeated exposure to the performance material. Participants also received researcher guidance during portions of the study, making the environment's effects difficult to distinguish from those of direct instruction. Analysis was descriptive, with no inferential statistics, and a single rater scored all performance assessments. Open-ended responses were used illustratively without formal qualitative coding, and accessibility findings reflect participant perceptions rather than measured changes in access or participation.

B. Future Work

Future work should address these limitations through pre-post or controlled designs. It should also examine longer-term skill retention and self-directed practice, compare instructional supports directly, and extend the digital twin approach to other mallet percussion instruments and XR music-learning environments. Overall, this work contributes an exploratory examination of learning, instructional support, and learner experiences within a digital twin musical environment, and of where virtual practice complements the physical instrument.

REFERENCES

- [1] National Center for Education Statistics, "Fast facts: Title I," U.S. Department of Education, Institute of Education Sciences, 2024, accessed: Aug. 20, 2026. [Online]. Available: <https://nces.ed.gov/fastfacts/display.asp?id=158>
- [2] N. Gibbs, "Inequitable Access in the Music Related Industries: Proposed Strategies and Directions," in *The Kyoto Conference on Arts, Media & Culture 2022: Official Conference Proceedings*. The International Academic Forum (IAFOR), 2022. [Online]. Available: https://papers.iafor.org/wp-content/uploads/papers/kamc2022/KAMC2022_65749.pdf
- [3] E. McGonigal, "A Comparison of Ratings of Bands from Title I Schools vs Non-Title I Schools at the Florida Music Performance Assessment," *Research Perspectives in Music Education*, vol. 21, no. 1, pp. 64–74, Dec. 2020.
- [4] S. Burak and O. Atabek, "Relationships Between Self-Regulated Learning, Big Five Personality Traits, and Amount of Instrument Practice," *Journal of Research in Music Education*, vol. 71, no. 2, pp. 208–228, Jul. 2023. [Online]. Available: <https://doi.org/10.1177/00224294221125183>
- [5] L. Turchet, R. Hamilton, and A. Çamci, "Music in Extended Realities," *IEEE Access*, vol. 9, pp. 15 810–15 832, 2021. [Online]. Available: <https://ieeexplore.ieee.org/document/9328440>

- [6] L. Turchet, "Musical Metaverse: vision, opportunities, and challenges," *Personal and Ubiquitous Computing*, pp. 1–17, Jan. 2023. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9837761/>
- [7] K. Kilteni, I. Bergstrom, and M. Slater, "Drumming in Immersive Virtual Reality: The Body Shapes the Way We Play," *IEEE transactions on visualization and computer graphics*, vol. 19, pp. 597–605, Apr. 2013.
- [8] B. Wu, X. Yu, and X. Gu, "Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis," *British Journal of Educational Technology*, vol. 51, no. 6, pp. 1991–2005, 2020. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1111/bjet.13023>
- [9] J. Jewel and T. Morelli, "Using Virtual Reality to Create an Inclusive Virtual Drumming Environment," in *Universal Access in Human-Computer Interaction. Theory, Methods and Tools*, M. Antona and C. Stephanidis, Eds. Cham: Springer International Publishing, 2019, pp. 569–577.
- [10] L. Turchet, N. Garau, and N. Conci, "Networked Musical XR: where's the limit? A preliminary investigation on the joint use of point clouds and low-latency audio communication," in *Proceedings of the 17th International Audio Mostly Conference*, ser. AM '22. New York, NY, USA: Association for Computing Machinery, Oct. 2022, pp. 226–230. [Online]. Available: <https://dl.acm.org/doi/10.1145/3561212.3561237>
- [11] S. Serafin, C. Erkut, J. Kojs, N. C. Nilsson, and R. Nordahl, "Virtual Reality Musical Instruments: State of the Art, Design Principles, and Future Directions," *Computer Music Journal*, vol. 40, no. 3, pp. 22–40, 2016. [Online]. Available: <https://www.jstor.org/stable/26777007>
- [12] J. L. Gomez-Sirvent, F. López de la Rosa, R. Sánchez-Reolid, R. Morales Herrera, and A. Fernández-Caballero, "Musical Instruments in Extended Reality: A Systematic Review," *International Journal of Human-Computer Interaction*, vol. 41, no. 16, pp. 10 143–10 162, Aug. 2025. [Online]. Available: <https://doi.org/10.1080/10447318.2024.2431352>
- [13] T. Loria, J. E. Teich, M. Pranjic, M. Tan, A. Huang, and M. H. Thaut, "The Impact of Limb Velocity Variability on Mallet Accuracy in Marimba Performance," *Journal of Motor Behavior*, vol. 54, no. 6, pp. 694–705, Nov. 2022. [Online]. Available: <https://doi.org/10.1080/00222895.2022.2069080>
- [14] J. Pinkl, J. Villegas, and M. Cohen, "Multimodal Drumming Education Tool in Mixed Reality," *Multimodal Technologies and Interaction*, vol. 8, no. 8, p. 70, Aug. 2024. [Online]. Available: <https://www.mdpi.com/2414-4088/8/8/70>
- [15] M. Banquero, G. Valdeolivas, and M.-C. Juan, "Enhancing Musical Learning Through Mixed Reality: A Case Study Using PocketDrum and Meta Quest 3 for Drum Practice," *Sensors*, vol. 25, no. 22, p. 6836, Jan. 2025. [Online]. Available: <https://www.mdpi.com/1424-8220/25/22/6836>
- [16] JamhackGames, "Percussive VR," 2020. [Online]. Available: <https://www.jamhack.games/>
- [17] E. Tanirgan, ElNeax, and C. Kasapoglu, "Paradiddle - Play drums without the limitations of the real world," Sep. 2022. [Online]. Available: <https://paradiddleapp.com/>
- [18] B. Zhong, S. Su, X. Liu, and Z. Zhan, "A literature review on the empirical studies of technology-based embodied learning," *Interactive Learning Environments*, vol. 31, no. 8, pp. 5180–5199, Nov. 2023. [Online]. Available: <https://doi.org/10.1080/10494820.2021.1999274>
- [19] S. Yu, Q. Liu, M. C. Johnson-Glenberg, M. Han, J. Ma, S. Ba, and L. Wu, "Promoting musical instrument learning in virtual reality environment: Effects of embodiment and visual cues," *Computers & Education*, vol. 198, p. 104764, Jun. 2023. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0360131523000416>
- [20] S. Willemsen, A. Horvath, and M. Nascimben, "DIGIDRUM - A HAPTIC-BASED VIRTUAL REALITY MUSICAL INSTRUMENT AND A CASE STUDY," Jun. 2020. [Online]. Available: <https://zenodo.org/record/4228322>
- [21] K.-H. Tai, J.-C. Hong, K.-F. Chen, and C.-L. Lin, "Practicing drum on VR to promote rhythm performance: Exploring the learning progress related to incremental belief of rhythm, gameplay anxiety, flow experience, and perceived learning value," *Entertainment Computing*, vol. 48, p. 100607, Jan. 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1875952123000629>
- [22] J. Pinkl and M. Cohen, "VR Drumming Pedagogy: Action Observation, Virtual Co-Embodiment, and Development of Drumming "Halvatar"," *Electronics*, vol. 12, no. 17, p. 3708, Jan. 2023. [Online]. Available: <https://www.mdpi.com/2079-9292/12/17/3708>
- [23] D. Hackl and C. Anthes, "HoloKeys - An Augmented Reality Application for Learning the Piano," 2017.
- [24] R. Guo, J. Cui, W. Zhao, S. Li, and A. Hao, "Hand-by-Hand Mentor: An AR based Training System for Piano Performance," in *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, Mar. 2021, pp. 436–437. [Online]. Available: <https://ieeexplore.ieee.org/document/9419239>
- [25] D. Johnson, D. Damian, and G. Tzanetakis, "Evaluating the effectiveness of mixed reality music instrument learning with the theremin," *Virtual Reality*, vol. 24, no. 2, pp. 303–317, Jun. 2020. [Online]. Available: <https://doi.org/10.1007/s10055-019-00388-8>
- [26] C. A. T. Fernandez, P. Paliyawan, C. C. Yin, and R. Thawonmas, "Piano learning application with feedback provided by an AR virtual character," in *2016 IEEE 5th Global Conference on Consumer Electronics*, Oct. 2016, pp. 1–2. [Online]. Available: <https://ieeexplore.ieee.org/document/7800380>
- [27] L. Turchet, C. Fischione, G. Essl, D. Keller, and M. Barthet, "Internet of Musical Things: Vision and Challenges," *IEEE Access*, vol. 6, pp. 61 994–62 017, 2018. [Online]. Available: <https://ieeexplore.ieee.org/document/8472621>
- [28] T. Hopkins, S. C. C. Weng, R. Vanukuru, E. A. Wenzel, A. Banic, M. D. Gross, and E. Y.-L. Do, "AR Drum Circle: Real-Time Collaborative Drumming in AR," *Frontiers in Virtual Reality*, vol. 3, Aug. 2022. [Online]. Available: <https://www.frontiersin.org/journals/virtual-reality/articles/10.3389/frvir.2022.847284/full>
- [29] T. Hopkins, R. Vanukuru, S. C. Chuan Weng, C. Tobin, A. Banic, M. D. Gross, and E. Yi-Luen Do, "Networking AI-Driven Virtual Musicians in Extended Reality," in *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. Sydney, Australia: IEEE, Oct. 2023, pp. 752–756. [Online]. Available: <https://ieeexplore.ieee.org/document/10322126/>
- [30] K. Schaeffer, "Key facts about public school teachers in the U.S." Sep. 2024. [Online]. Available: <https://www.pewresearch.org/short-reads/2024/09/24/key-facts-about-public-school-teachers-in-the-u-s/>
- [31] League of American Orchestras, "Racial/Ethnic and Gender Diversity in the Orchestra Field in 2023," League of American Orchestras, Tech. Rep., Jun. 2023. [Online]. Available: <https://americanorchestras.org/wp-content/uploads/2023/06/Racial-Ethnic-and-Gender-Diversity-in-the-Orchestra-Field-in-2023.pdf>
- [32] J. La Favre, "Marimba bar and resonator data," 2007. [Online]. Available: <https://www.lafavre.us/marimba-bars.htm>
- [33] Z. Ren, R. Mehra, J. Copsky, and M. C. Lin, "Tabletop Ensemble: touch-enabled virtual percussion instruments," in *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*. Costa Mesa California: ACM, Mar. 2012, pp. 7–14. [Online]. Available: <https://dl.acm.org/doi/10.1145/2159616.2159618>
- [34] Y. Gao and C. Spence, "Enhancing Presence, Immersion, and Interaction in Multisensory Experiences Through Touch and Haptic Feedback," *Virtual Worlds*, vol. 4, no. 1, p. 3, Jan. 2025. [Online]. Available: <https://www.mdpi.com/2813-2084/4/1/3>
- [35] P. Hanappe, "FluidSynth," 2000. [Online]. Available: <https://www.fluidsynth.org>
- [36] Paxsteller, "Maestro MIDI Player Toolkit," 2024. [Online]. Available: <https://paxstellar.fr/>
- [37] Daindune, "MalletsSectionGM SoundFont," Jan. 2023. [Online]. Available: <https://musical-artifacts.com/artifacts/2650>
- [38] RhythmBandInstrumentsLLC, "Boomwhackers Tuned Percussion Tubes," Jul. 2009. [Online]. Available: <https://boomwhackers.com/pages/about-boomwhackers>
- [39] G. Xia, D. Chin, Y. Zhang, T. Zhang, and J. Zhao, "Interactive Rainbow Score: A Visual-centered Multimodal Flute Tutoring System." Zenodo, Jun. 2020. [Online]. Available: <https://zenodo.org/record/4813324>
- [40] A. N. Grey, "Rote Instruction in Secondary Instrumental Music Classrooms: A Review of the Literature," *Update: Applications of Research in Music Education*, vol. 39, no. 1, pp. 59–69, Oct. 2020. [Online]. Available: <https://doi.org/10.1177/8755123320909149>
- [41] D. Workman, *The Percussionists' Guide to Injury Treatment and Prevention: The Answer Guide for Drummers in Pain*. Taylor & Francis, 2006.