Sound Blocks VR: An Accessible Virtual Reality Musical Instrument

Marta Gioiosa[1,2], Federico Avanzini[1], Luca Andrea Ludovico[1(✉)], Susanna Brambilla[2], and Laura Ripamonti[2]

[1] Laboratory of Music Informatics (LIM), Department of Computer Science, University of Milan, Milan, Italy
marta.gioiosa@studenti.unimi.it,
{federico.avanzini,luca.ludovico}@unimi.it

[2] Playlab fOr inNovation in Games (PONG), Department of Computer Science, University of Milan, Milan, Italy
{susanna.brambilla,laura.ripamonti}@unimi.it
https://www.lim.di.unimi.it, https://pong.di.unimi.it/

Abstract. In recent years, immersive technologies such as augmented and virtual reality have gained significant popularity and found applications across various research domains. This paper focuses on the adoption of such technologies for musical purposes, particularly in enhancing accessibility in musical performance. In the domain of Digital Musical Instruments, extensive efforts have been made to develop accessible designs and interfaces for music creation. Concerning Virtual Reality Musical Instruments (VRMIs), one notable challenge is the discomfort stemming from using controllers to interact with objects in the virtual world, specifically for people with a physical disability. In fact, the design of these applications often incorporates objects that require precise control. The proposed solution is an accessible VRMI, called *Sound Blocks VR*, which accommodates both controller-based and hand-tracking interactions to address the needs of individuals with physical disabilities (especially motor-impaired users) or, in general, who feel uncomfortable using controllers. Where available, the passthrough feature can also be exploited. The usability of *Sound Blocks VR* was evaluated using the System Usability Scale through interviews with musicians and nonmusicians. The average usability score computed on 17 participants is 87.94 out of 100, which indicates that participants found the experience intuitive and user-friendly. Moreover, some qualitative questions were posed to investigate specific aspects of the experience design.

Keywords: Digital Musical Instruments · Virtual Reality · Accessibility · Motor Impairment

1 Introduction

Digital Musical Instruments (DMIs) were developed to enable users to create music by providing input that is translated into musical parameters and processed by a digital synthesizer. In this field, considerable work has been dedicated to accessibility, leading to the creation of new interfaces known as Accessible Digital Music Instruments

H. Plácido da Silva and P. Cipresso (Eds.): CHIRA 2024, CCIS 2371, pp. 316–330, 2025.
https://doi.org/10.1007/978-3-031-83845-3_19

(ADMIs) [11]. Many ADMIs come in the form of Tangible User Interfaces (TUIs), namely interfaces that allow users to interact with digital music through physical objects and spaces [21]. Music TUIs integrate physical objects into the digital interaction, making the user experience more intuitive and tactile.

Other interface categories have been developed, such as eye tracking, head movement, mouth aperture, or other movements captured by various sensors to translate body movements into musical parameters [6]. An example is the *Tongue-Controlled Electro-Musical Instrument* [16], an ADMI that exclusively uses the tongue (mouth) as the physical channel, designed mainly for people who cannot move their arms. These innovations aim to make music creation more inclusive and accessible to individuals with diverse abilities and preferences.

In recent years, new technologies have been developed that introduce novel techniques for creating music. For example, augmented reality (AR) and virtual reality (VR) let users have immersive experiences, enabling them to explore virtual worlds and also perform actions that are impossible in the physical world. These technologies have garnered significant interest in the field of music computing, leading to the development of numerous DMIs that users can play in an immersive environment. These instruments are known as Virtual Reality Musical Instruments (VRMIs) [20].

In analyzing some examples of VRMIs, as reported in the next section, we observed that input strategies can require a high level of precision. For instance, when using controllers, the user must point at a specific object and press a button to trigger events. In addition, raycasting is often used to set parameters or interact with virtual objects. Raycasting is a computational technique that is used in virtual environments to determine the visibility and interaction of objects within a 3D space from a specific point of view. Unfortunately, raycasting relies on precise motor feedback. Users with motor impairments might struggle with environments where fine details and exact positioning are crucial. In addition, low contrast and poor visibility of objects can make it difficult for visually impaired users to navigate and interact effectively. Finally, navigating a raycast-based environment can require precise control and quick reflexes, posing challenges for users with limited motor skills or coordination problems. Please note that many raycast interactions are designed for mouse or touch controls, which might not be suitable for all users, particularly those relying on adaptive devices. In conclusion, approaches that require high precision can be very uncomfortable or even impossible for people with physical disabilities, such as those who have difficulty performing precise movements due to dystonia or rigid muscle movements.

The analysis carried out so far highlights that, unlike in DMIs, the accessibility theme is not adequately considered in the design and development of VRMIs. In contrast, this work focuses on enabling music performance and expression for people with physical disabilities that allow them to move only their upper limbs and produce imprecise movements. Our target users should be able to play the instrument in a virtual world by touching virtual objects with their hands or controllers. To this end, we have designed and released an ADMI called *Sound Blocks VR*. Developed in Unity, it adopts several libraries to manage virtual components, support collaborative performances in augmented or virtual reality, and control a synthesizer via the Musical Instrument Digital Interface (MIDI) protocol.[1] Future plans include integrating this VRMI into a VR platform hosting various VRMIs, thus enhancing versatility and accessibility in music creation.

The remainder of the paper is structured as follows. In Sect. 2, we will present other experiences, products, and scientific works that are relevant to our field; in Sect. 3, we will focus on *Sound Blocks VR*, namely the ADMI that we have developed to address the theme of accessibility in music creation; in Sect. 4, we will describe the early experimentation conducted with 17 subjects in order to evaluate the level of efficacy and user acceptance; in Sect. 5, we will discuss the results achieved; finally, In Sect. 6, we will draw the conclusions and pave the way for future developments.

2 Related Works

2.1 Accessible Digital Musical Instruments

Accessible Digital Musical Instruments pave the way toward new opportunities for inclusion by accommodating bodily differences and valuing multiple ways of perceiving music. The analysis carried out by Frid in 2019 on available ADMIs [12] underlined that most ADMIs were tangible or physical controllers.

Tangible music user interfaces (TUIs) definitely play an important role in designing new ways of controlling musical parameters by providing physical and intuitive methods to manipulate sound. This approach makes music creation more accessible to people with diverse abilities [2].

One of the first examples of TUI was the *reacTable* [14], a collaborative electro-acoustic musical instrument. The instrument hardware consisted of a translucent round table with a video camera positioned beneath it, continuously analyzing the table's surface. The tangible objects on the *reacTable* represented components of a classic modular synthesizer. Users interact with these objects by moving them and changing their positions, orientations, or faces. These interactions modify the primary characteristics of the sounds produced, allowing for dynamic and intuitive music creation. The *reacTable*, now out of productions, had not only a relevant scientific impact but also practical and artistic applications, being used by important performers and DJs in their shows.

Another music TUI is *Kibo* [1], a MIDI controller based on a simplified tangible interface composed of eight geometric extractable solids and a knob to control MIDI parameters. As with many ADMIs, *Kibo* can have therapeutic purposes when used in a rehabilitation context, addressing people with cognitive or physical impairments. The physical identity attributed to music parameters simplifies the instrument's use for blind or visually impaired (BVI) people and those with conditions like autism. The unique shape of each tangible object makes it easy for BVI users to trigger music events. This ADMI also serves as a compensatory enabling technology, thanks to its fine-sensing capabilities. Being able to detect even small movements, this device can also be used profitably in elderly rehabilitation [15].

Other types of interface have been developed, such as those that utilize eye tracking or various sensors to translate movements into musical parameters. These innovations

[1] https://midi.org.

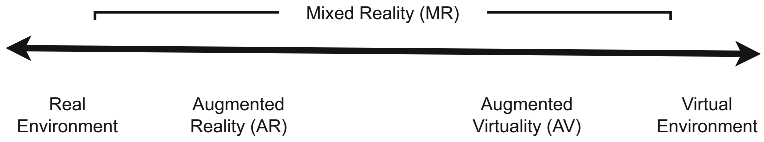

Fig. 1. The Milgram-Kishino Virtuality Continuum. Image adapted from [23].

aim to make music creation more inclusive and accessible. An example is *EyeHarp* [24], an ADMI based on eye tracking that allows users to perform and compose music by controlling sound settings and musical events using eye movements. The pupil is tracked by a camera that detects the fixation point with specialized software, allowing for real-time performances or musical compositions through a step sequencer.

The research field of eye tracking applied to music and sound parameter control has been widely explored at the Laboratory of Music Informatics, University of Milan. The results included design guidelines for gaze-based DMIs [7] and the release of several computer-based ADMIs, such as Netychords [8], *Kiroll* [9], and DJeye [3]. These musical instruments, even if usable by any kind of user, turned out to be particularly suitable for people with quadriplegia.

2.2 Virtual Reality Musical Instruments

Physically disabled people can use novel technological tools as extensions of themselves. One of the possibilities offered by enabling technologies is to play music with augmented instruments. In this sense, the widespread use of VR in both professional and recreational settings is bringing significant advances. Initially, audio was used in VR scenarios to enhance the sense of presence and spatial impression in the virtual experience. Then, the novel concept of Musical Extended Reality (XR) [23] emerged and gained traction, leading to the creation of many VR/AR musical instruments.

As shown in Fig. 1, XR can be considered an umbrella term that encompasses a range of systems and experiences commonly referred to as Virtual Reality, Augmented Reality, Augmented Virtuality, and Mixed Reality. Musical XR projects can take a diverse range of forms, including VRMIs, immersive concert experiences, generative musical systems, and gamified musical environments. Playing virtual objects, users can experience new dimensions of interaction and creativity in music making. Please note that, even if in this work we are addressing *accessible* musical instruments and we conducted tests with impaired users, these interfaces are suitable for many other applications, including music education in young learners and new forms of expressivity for musicians and nonmusicians with no impairment.

A notable example of VRMI is *ChromaChord* [10], which uses a combination of *Oculus Rift DK2* to immerse the performer in the virtual world and *Leap Motion* mounted on the headset for infrared hand tracking [22]. To make music, users must place their hands vertically in front of them and "touch" the keys, which are arranged for ease of use and the ability to play multiple notes simultaneously. Additionally, touching virtual objects with the hand allows users to change various sound parameters.

Another example of VRMI is *Cirque des Bouteille* [26], where the user blows into a microphone and the stream of air is recreated in the virtual environment, directed toward the virtual bottles that the user is pointing to with fingers, thus creating the corresponding sound. Fingers are tracked using *Leap Motion* and the user wears an *Oculus Rift* to become immersed in the scene.

Another relevant work is *Crosscale* [5], a 3D virtual instrument where objects serve as metaphors for note keys organized in multiple lines, forming a playable spatial instrument. The player can perform sequences of notes and chords across the scale using short gestures that minimize jump distances. The players can use both hands to interact with the system, driving two independent virtual cursors using *Razer Hydra* controllers and several buttons. The system uses *Oculus Rift* for immersion in the virtual world. To play a note, users have to hover the 3D cursor over the corresponding 3D object and press the trigger. To play a continuous sequence, users can trigger a note and then move along a path to the subsequent notes.

Musical instruments in AR rather than VR are relevant to our interest as well. We will address the strengths and weaknesses of this approach in Sect. 5.2. An example of an AR musical instrument is *Augmented Groove* [19], where users can play together with or without traditional musical instruments. The players wear an HMD that shows some cards that they can manipulate to control musical parameters. Encouraging several people to jam together, *Augmented Groove* can be considered a collaborative musical environment.

Unfortunately, the adoption of VRMIs is hampered by problems such as addressing a niche of users, requiring specific technologies, and dealing with typically conservative environments such as that of classical music. Moreover, these instruments lack a tradition of practice and a historical repertoire. As such, VRMIs have not become firmly established for the majority of performers and composers, yet.

3 Sound Blocks VR

Sound Blocks VR is a VRMI that was specifically designed to address motor-impaired users. In general terms, it can be considered a MIDI controller coupled with an embedded MIDI synthesizer.

MIDI, which stands for Musical Instrument Digital Interface, is a standard protocol that allows electronic musical instruments, computers, and other equipment to communicate, control, and synchronize with each other by sending digital instructions related to musical performance and sound production.

Sound Blocks VR is freely available in the form of an Android application package (Android Package Kit, APK) at https://drive.google.com/file/d/1TKmtcN3NYhjutIRSk2Asxbth55V1UbWN/view?usp=share_link. The code falls under the definition of Free and Open Source Software (FOSS) and is distributed under a CC BY-NC license. The source code is available at https://github.com/LIMUNIMI/Sound-Blocks-VR.

In the rest of the section, we will describe the technological aspects and design choices of *Sound Blocks VR*.

3.1 Technological Aspects

Sound Blocks VR uses the *Meta Quest* technology to provide immersion in the virtual world. It was developed in *Unity* adopting the *Meta XR All-in-One SDK* to support VR. This library enables passthrough, supports hand and controller tracking, includes custom interactable objects, and recognizes specific gestures. *Meta Quest* uses its cameras integrated in the HMD along with advanced computer vision algorithms to track hand movements. Controller tracking employs sensors within the controllers, infrared LEDs that are visible to the cameras, and computational algorithms to process sensor data and determine orientation and position.

As mentioned above, *Sound Blocks VR* is a MIDI controller with sound synthesis capabilities. MIDI was chosen for several reasons. First, it is a well-established protocol in the field of sound and music computing. As such, it is largely supported by dedicated hardware and software, including ad-hoc libraries in many programming languages. Moreover, it is an easy but effective protocol to manage music generation, at least according to Western music theory and the 12-tone equal temperament (12 TET). Finally, the quality of sound production and the destination of MIDI messages can be configured. In this prototype, everything is managed internally, using an embedded MIDI synthesizer and its built-in timbres, but, in the future, the controller will be able to send messages to an external MIDI system, providing, e.g., better sound quality thanks to dedicated sound-generation devices or extended functions depending on how the setup will manage the messages coming from the virtual instrument.

To integrate MIDI into *Unity*, we used the free version of the *MIDI Maestro Toolkit*.[2] This library provides classes to control MIDI streams and send MIDI messages. Interestingly, the *MIDI Maestro Toolkit* allows the association of different patches on a single channel, a feature that does not adhere to the MIDI standard. In order to maintain compliance, a restriction was implemented to assign a single patch to each channel. Specifically, we adopted the *General MIDI SoundFont*, with channel 10 dedicated to percussive sounds.

The sound generation relies on *FluidSynth*,[3] a real-time software synthesizer built on the *SoundFont 2* specifications and implementing wave-table synthesis.

3.2 Design

During the entire design process, the goal was to develop a simplified interaction method that enhances both the performer's expressivity and accessibility. To establish a starting point for development and guide the design phase, a team member with a motor impairment, who can move their upper limbs without imprecise movements, was involved. This individual found the use of controllers particularly uncomfortable, especially when it came to pressing physical buttons to interact with the virtual environment. Pointing to a specific location while pressing buttons simultaneously also proved challenging. These difficulties highlighted the need for an alternative approach, leading us to prioritize hand tracking over controllers, that better accommodates users with motor impairments.

[2] https://paxstellar.fr.

[3] https://www.fluidsynth.org.

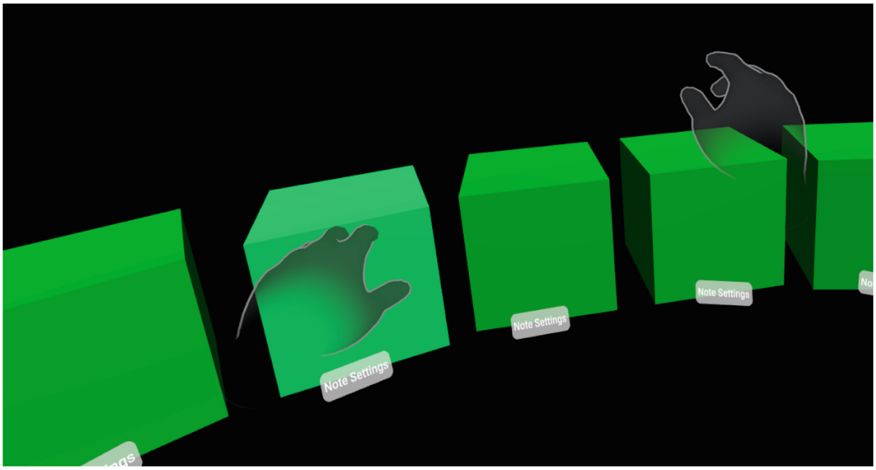

Fig. 2. Playing a melody in *Sound Blocks VR* by touching sound cubes.

The uniqueness of our work lies in allowing users to interact directly with virtual objects with their hands. Additionally, users benefit from an interface that exclusively utilizes virtual buttons, allowing the system to function without requiring precise movements. This design enables interaction with their hand using hand-tracking technology, making it more accessible for users.

Hand-tracking was chosen to allow direct interaction with virtual objects, as shown in Fig. 2. The use of raycasting was avoided as it requires a high level of precision when pointing at objects, potentially posing critical issues to some categories of impaired users.

Notes are represented in the immersive environment in the form of sound cubes sensitive to collisions. Players can interact in the immersive environment using either their hands or controllers. In both modalities, it is possible to trigger sounds by touching the virtual blocks and manage settings by "poking" the buttons.

Here is a simple explanation of how the MIDI controller currently works. When a collision between the virtual-hand mesh[4] and the object mesh is detected, the system sends a Note-On MIDI message that triggers the sound; when the collision stops, namely the hand exits the object box collider, the system sends the corresponding Note-Off message. These commands are sent to the internal synth for sound generation.

Users are also allowed to customize the experience. An *ad hoc* canvas lets players add single objects or presets for musical scales. Users can add as many sound cubes as they want and rearrange each part of the interface in the virtual environment. For example, it is possible to change the position of a sound cube simply by grabbing it. The system recognizes when users close their hands, allowing them to move the objects. The same action can be performed through the game controllers by touching virtual

[4] Please note that, in our approach, the virtual hand positions can be reconstructed from the detection of either the real hands, captured by cameras, or the game controllers held by the player.

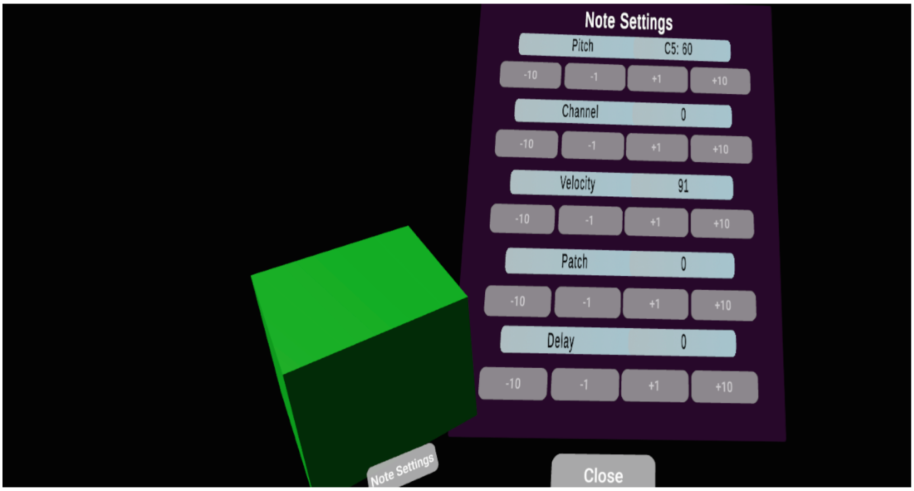

Fig. 3. Note Settings.

objects and pressing the grab button to move them. This feature helps to make performance more comfortable and user-tailored, an aspect particularly relevant in the field of accessibility.

Each sound cube has a button on its bottom that opens a canvas to set various music parameters (see Fig. 3). Here, users can adjust the pitch associated with the object, the MIDI channel,[5] velocity, and patch, and set a delay. Users can change values simply by using buttons; even if sliders or other types of controls could be more intuitive and effective, the adoption of buttons solves the issue of imprecise movements. This type of interaction is technically known as "poke interaction". Canvases are always oriented toward the user's position. Players can move the Note Settings canvas by grabbing the sound cube and moving it.

Users can also delete virtual objects by dragging them to a trash bin located on the left in the virtual environment.

4 Evaluation

To evaluate this project, 17 subjects aged 18 to 50, both musicians and nonmusicians, were involved. The tests were carried out under the supervision of an expert between June and July 2024 at the Playlab fOr inNovation in Games (PONG), Department of Computer Science, University of Milan.

[5] Please note that, in the MIDI protocol, it is possible to associate different program numbers (i.e. different patches or timbres) to different channels, thus turning the instrument into a multi-timbral one.

4.1 Methodology

Each test session was structured into two phases.

In the former phase, participants were asked to try *Sound Blocks VR* and test its main features. Specifically, they were instructed to customize the layout of the sound cubes, adjust musical parameters, and play the instrument. During this step, some users started the test using controllers and subsequently transitioned to the hand-tracking mode, while others began with the hand-tracking and later switched to the controller-tracking mode. Users were also asked to provide feedback during the experience.

The second phase of the test consisted of answering a questionnaire. In the first part of the survey, participants were asked to provide anonymous profile information, such as their age, level of musical knowledge, and whether they had prior experience with VR. Following the initial questions, participants were presented with specific inquiries regarding their experience. Participants had to express their preference for the use of controllers or hand tracking.

The following part of the questionnaire aimed to assess the effectiveness of canvas design, accessibility, and expressiveness of the performance using the System Usability Scale (SUS). Proposed by Brooke in 1996 [4], SUS is a standardized tool to measure the usability of various systems in different fields. In this context, the aim was to estimate how simple interaction was perceived, how well features had been organized, and how users evaluated the experience in general, thus providing feedback to improve the system's quality.

SUS consists of 10 statements, forming a 10-item scale that offers a global view of subjective usability assessments. For each statement, respondents indicate their degree of agreement or disagreement on a 5-point Likert scale, where 1 corresponds to "Strongly disagree" and 5 to "Strongly agree". SUS statements are:

1. I think that I would like to use this system frequently;
2. I found the system unnecessarily complex;
3. I thought the system was easy to use;
4. I think that I would need the support of a technical person to be able to use this system;
5. I found the various functions in this system were well integrated;
6. I thought there was too much inconsistency in this system;
7. I would imagine that most people would learn to use this system very quickly;
8. I found the system very cumbersome to use;
9. I felt very confident using the system;
10. I needed to learn a lot of things before I could get going with this system.

SUS statements are intentionally very general. They can provide an idea of the usability perceived by a system's user but are not intended to investigate specific goals or functions of the system. For this reason, we also posed additional qualitative questions that will be listed in the next section, together with the results.

4.2 Results

The questionnaire revealed that 76.57% of the sample had previous knowledge of music. This aspect must be taken into consideration in the analysis of the results but

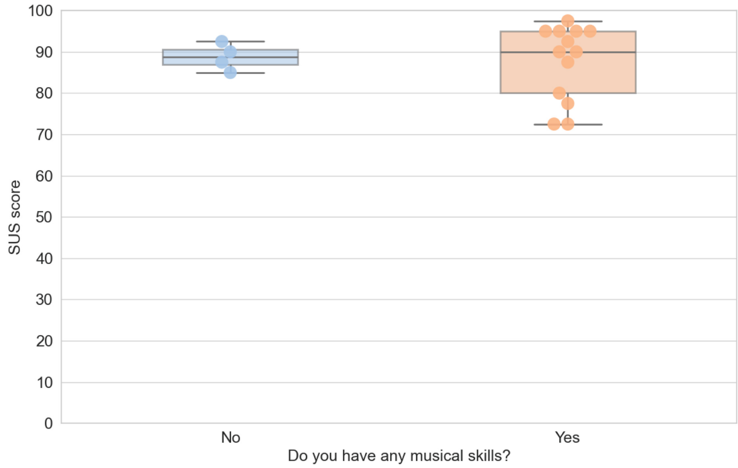

Fig. 4. Comparison of SUS scores between nonmusicians (left) and musicians (right). The swarm plot displays individual scores, highlighting the distribution of ratings within each group. The overlaid box plots provide a summary of descriptive statistics, illustrating the median, quartiles, and variability of scores for each category.

is not in contrast with one of the intended uses of *Sound Blocks VR*, which can be played as a virtual instrument also by musicians. Furthermore, 94.1% of the sample declared that they had already experienced immersive environments.

The SUS score was evaluated for each participant, with the lowest score being 72.5 out of 100 and the highest being 97.5 out of 100. In the SUS framework, the average score of the 17 participants is 87.94 out of 100, which translates to an A grade (excellent).

Since the musical experience involved both musicians and nonmusicians, we calculated the averages of SUS scores for the two groups separately. For the former group, the average score was 87.69 out of 100, translating to an A grade (excellent). For the latter, the average score was 88.75 out of 100, which also translated to an A grade (excellent). The two distributions are shown in Fig. 4.

This is the list of qualitative questions and their results:

– "Did you prefer the interaction with hand-tracking or with controllers?"
  Hand tracking: 82.4%, Controllers: 17.6%;
– "Do you think that the use of hand-tracking made the performance more expressive?"
  Not at all: 0%, Little: 5.9%, Moderately: 17.6%, Considerably: 52.9%, Extremely: 23.5%;
– "Do you think that the use of hand-tracking made the performance more accessible?"
  Not at all: 0%, Little: 0%, Moderately: 47.1%, Considerably: 35.3%, Extremely: 17.6%;
– "Did you find the design of the canvas comfortable for setting MIDI parameters of individual objects (pitch, channel, patch, ...)?"

Not at all: 0%, Little: 11.8%, Moderately: 17.6%, Considerably: 64.6%, Extremely: 5.9%;
- "Did you find any difficulties in playing and/or forming the instrument?" at all: 17.6%, Little: 47.1%, Moderately: 9.4%, Considerably: 5.9%, Extremely: 0%;
- "Did you perceive any latency between touching the virtual object and the emission of the related sound?" at all: 52.9%, Little: 23.5%, Moderately: 11.8%, Considerably: 5.9%, Extremely: 5.9%.

During the interview, participants were also asked to provide personal feedback and advice to improve the usability of the instrument. Some of these responses were highly relevant, while others did not align with the instrument's objectives or were purely aesthetic. Below, we report relevant feedback:

- Add vibrotactile feedback to controllers;
- Provide the option to activate/deactivate the grab gesture;
- Differentiate cubes by indicating pitch with labels and colors according to chromesthesia;
- Incorporate control over velocity intensity based on how strongly the user touches the cube;
- Prioritize open settings in the foreground.

5 Discussion

5.1 Comments on Test Results

In evaluating the results, we have to underline that the tests were conducted in an academic setting where the participants were mostly musicians. Most of them had previous experience with VR, suggesting that this technology has gained success in recent years.

In this initial phase, our goal was to evaluate the general usability of the experience and collect feedback on how to improve the user experience. Unfortunately, none of the participants were impaired users. This problem will be addressed in the near future by involving institutions and caregivers.

All participants enjoyed the proposed activity. In general, users expressed their preference for the hand-tracking experience, which offered more freedom in gestures and represented a novel form of control. However, a few participants preferred the use of controllers due to higher precision and lower latency, which are particularly relevant aspects in rhythmic performances. In addition, some categories of players are accustomed to holding tools in their hands, such as drumsticks and mallets.

Most participants found the interface design intuitive. After a few minutes of use, they could profitably exploit its features, complete the assigned tasks, and create a short musical performance.

The results on the perception of latency varied from user to user, as shown in Fig. 5. In this test, latency was evaluated as a perceptual factor, meaning that it largely depended on the user's habits. For example, musicians perceived a higher level of latency than nonmusicians, and even more so for drummers who are accustomed to precise note emission. Also the self-assessed familiarity with VR influenced these results.

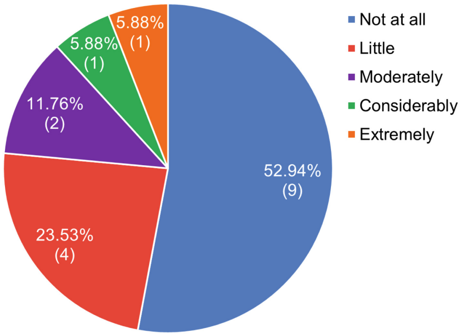

Fig. 5. Answers to the question: "Did you perceive any latency between touching the virtual object and the emission of the related sound?"

5.2 Transitioning From VR to AR

Thanks to the passthrough feature available on some headsets, e.g. *Meta Quest 3*, it is possible to use *Sound Blocks VR* in AR rather than VR. As a result, the described one-player experience can turn into a collaborative performance involving multiple persons playing either traditional or virtual instruments.

This approach offers several advantages over pure VR. First, the integration of one or more ADMIs with other musical instruments encourages communication and collaboration between musicians. Based on the universality of music, the positive effects range from the promotion of social interaction for people with communication impairments [25] to the exploration of different levels of pre-intentional and intentional communication development in musical interaction with children with severe and multiple disabilities [18]. The goal is to move from performance to participatory-oriented musical expression [17].

Second, the adoption of an AR approach can reduce the risks of escapism, prevent the loss of connection with reality, and alleviate the sense of isolation that is often recognized as one of the most critical issues in the adoption of VR technologies, especially in educational and social settings [13].

Finally, if used in a social context, *Sound Blocks VR* is not only an enabling technology for impaired users. Rather, AR can turn this ADMI into an opportunity to enrich the ensemble with new timbral possibilities and extended expressivity. The impaired musician can participate and even become the soloist at the center of the performance.

Unfortunately, in the transition from VR to AR, there are some issues that we need to consider. Latency in the perception of other players' movements and performance is a crucial factor in making music together. A tolerance threshold often mentioned for delays on digital audio workstations is 5 ms. In devices such as *Meta Quest 3*, external images are acquired through cameras and projected onto headset lenses; this process introduces additional latency. Since there are many AR gaming applications based on timely interaction between generated images and the real world, the aspect of latency introduced by cameras is well studied and we can expect better performances in future products. Another issue concerns the acquisition of the external audio environment, which has to be reproduced in real time to obtain a satisfying experience. These new latency issues add to those mentioned in Sect. 5.1. Finally, the quality of microphones and built-in speakers must be adequate. Specialized forums suggest a set of wired headphones or a purpose-built set of gaming earbuds connected via a USB-C dongle for a low-latency link.

Even if the described test campaign focused on VR and, consequently, did not consider the possibility of making music together, this scenario will be explored in our next steps.

6 Conclusion and Future Work

This paper presented *Sound Blocks VR*, an ADMI for immersive environments that is playable by musicians and nonmusicians, with or without impairments. Even if agnostic about impairments, *Sound Blocks VR* has been designed to address users with motor impairments. In an early test campaign involving 17 participants, heterogeneous in their music skills, the system received a SUS score indicating a high level of usability. The difference in average scores between musicians and non-musicians was not significant.

Some limitations need to be considered, which could influence the results, particularly regarding the characteristics of the participant group selected for the test. First, none of the participants were motor-impaired, which prevented us from gathering specific insights from target users on how to improve the system's accessibility. We have not been able to involve impaired users so far and this is undoubtedly a flaw in the early experimental activity presented in this work. Another limitation is that most participants, except one, had prior experience with VR, which could have impacted their familiarity and interaction with the system. Additionally, the majority of participants were musicians, which may have influenced their understanding and evaluation of the musical aspects of the system.

In the near future, we are planning to extend the testing to different categories of impaired musicians and nonmusicians, assigning specific tasks and assessing impaired users' ability to perform them.

The tests highlighted a potential issue, namely the latency of the system, which can pose critical problems in professional music applications and ensemble music. This factor depends not only on the developed software but also on the characteristics of the hardware device that runs it. Anyway, industrial interest and technological advances in the field of VR headsets should limit hardware-related latency effects in the near future. In the future, more tests will be conducted to evaluate latency from a technical perspective. This will involve analyzing the various components of the sound chain and attempting to reduce the multiple sources of delay.

We are planning to release a new version of *Sound Blocks VR* able to communicate with an external MIDI system. In this way, this virtual instrument will be integrated into a larger platform capable of collecting various VRMIs and delegating sound synthesis to external synthesizers. Another option is to use the MIDI messages generated in the immersive environment as control signals for extra-musical applications, e.g., in an integrated gamification environment aiming at music education.

Finally, a future goal will be to evaluate the performance from a musical perspective. More tests will be conducted to assess users' capability to carry out specific music tasks.

Shining a spotlight on accessibility, tests will be performed with people with physical disabilities to evaluate their ability to perform specifically designed tasks.

References

1. Amico, M.D., Ludovico, L.A.: Kibo: A MIDI controller with a tangible user interface for music education. In: Lane, H., Uhomoibhi, J., Zvacek, S. (eds.) Proceedings of the 12th International Conference on Computer Supported Education (CSEDU 2020), vol. 1, pp. 613–619. CSEDU, SCITEPRESS - Science and Technology Publications, Lda., Setúbal (2020). https://doi.org/10.5220/0009805206130619
2. Baratè, A., Korsten, H., Ludovico, L.A., Oriolo, E.: Music tangible user interfaces and vulnerable users: state of the art and experimentation. In: Constantine, L., Holzinger, A., Silva, H.P., Vanderdonckt, J. (eds.) CHIRA 2021. CCIS, vol. 1882, pp. 1–25. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-41962-1_1
3. Bottarelli, F., Davanzo, N., Presti, G., Avanzini, F., et al.: Djeye: towards an accessible gaze-based musical interface for quadriplegic DJs. In: Proceedings of the Sound and Music Computing Conferences, pp. 117–124. Sound and Music Computing Network (2023)
4. Brooke, J.: SUS: A 'quick and dirty' usability scale (1996)
5. Cabral, M., et al.: Crosscale: a 3D virtual musical instrument interface. In: 2015 IEEE Symposium on 3D User Interfaces (3DUI), pp. 199–200. IEEE (2015). https://doi.org/10.1109/3DUI.2015.7131769
6. Davanzo, N., Avanzini, F.: Hands-free accessible digital musical instruments: conceptual framework, challenges, and perspectives. IEEE Access **8**, 163975–163995 (2020)
7. Davanzo, N., Avanzini, F., et al.: Design concepts for gaze-based digital musical instruments. In: Proceedings of the 19th Sound and Music Computing Conference, pp. 472–478. SMC Network (2022)
8. Davanzo, N., De Filippis, M., Avanzini, F., et al.: Netychords: an accessible digital musical instrument for playing chords using gaze and head movements. In: CHIRA, pp. 202–209 (2021)
9. Davanzo, N., Valente, L., Ludovico, L.A., Avanzini, F.: Kiroll: a gaze-based instrument for quadriplegic musicians based on the context-switching paradigm. In: Proceedings of the 18th International Audio Mostly Conference, pp. 59–62 (2023)
10. Fillwalk, J.: ChromaChord: a virtual musical instrument. In: 2015 IEEE Symposium on 3D User Interfaces (3DUI), pp. 201–202. IEEE (2015). https://doi.org/10.1109/3DUI.2015.7131770
11. Frid, E.: Accessible digital musical instruments: a survey of inclusive instruments presented at the NIME, SMC and ICMC conferences. In: International Computer Music Conference 2018, pp. 53–59. The International Computer Music Association (2018)
12. Frid, E.: Accessible digital musical instruments-a review of musical interfaces in inclusive music practice. Multimodal Technol. Interact. **3**(3), 57 (2019). https://doi.org/10.3390/mti3030057
13. Hayes, A.: Will virtual reality connect or isolate students? In: Bridging human Intelligence and Artificial Intelligence, pp. 139–150. Springer (2022)
14. Jordà, S., Geiger, G., Alonso, M., Kaltenbrunner, M.: The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. In: Proceedings of the 1st International Conference on Tangible and Embedded Interaction, pp. 139–146 (2007)
15. Korsten, H., Baratè, A., Ludovico, L.A.: Design, implementation, and early experimentation of a music tangible user interface for elderly people rehabilitation. In: Constantine, L., Holzinger, A., Silva, H.P., Vanderdonckt, J. (eds.) CHIRA 2021. CCIS, vol. 1882, pp. 108–129. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-41962-1_6

16. Niikawa, T.: Tongue-controlled electro-musical instrument. In: Proceedings of the 18th International Congress on Acoustics, Kyoto, 2004, pp. 1905–1908 (2004)
17. Nilsson, B.: Performance and participation: a qualitative study of music education practices in digitally-based musicking with young people with physical impairments. Approaches: music Ther. Spec. Music Educ. **6**(1), 19–27 (2014)
18. Perry, M.M.R.: Relating improvisational music therapy with severely and multiply disabled children to communication development. J. Music Ther. **40**(3), 227–246 (2003). https://doi.org/10.1093/jmt/40.3.227
19. Poupyrev, I., et al.: Augmented groove: collaborative jamming in augmented reality. In: ACM SIGGRAPH 2000 Conference Abstracts and Applications, vol. 17, p. 77 (2000)
20. Serafin, S., Erkut, C., Kojs, J., Nilsson, N.C., Nordahl, R.: Virtual reality musical instruments: state of the art, design principles, and future directions. Comput. Music. J. **40**(3), 22–40 (2016)
21. Shaer, O., Hornecker, E., et al.: Tangible user interfaces: past, present, and future directions. Found. Trends® Hum. Comput. Interact. **3**(1–2), 4–137 (2010)
22. Silva, E.S., de Abreu, J.A.O., de Almeida, J.H.P., Teichrieb, V., Ramalho, G.L.: A preliminary evaluation of the leap motion sensor as controller of new digital musical instruments, pp. 59–70. Recife, Brasil (2013)
23. Turchet, L., Hamilton, R., Çamci, A.: Music in extended realities. IEEE Access **9**, 15810–15832 (2021). https://doi.org/10.1109/ACCESS.2021.3052931
24. Vamvakousis, Z., Ramirez, R.: The EyeHarp: an eye-tracking-based musical instrument. In: Proceedings of the 8th Sound and Music Computing Conference, SMC 2011 (2011)
25. Zeedyk, M.S.: Promoting Social Interaction for Individuals with Communicative Impairments: Making Contact. Jessica Kingsley Publishers (2008)
26. Zielasko, D., et al.: Cirque des bouteilles: the art of blowing on bottles. In: 2015 IEEE Symposium on 3D User Interfaces (3DUI), pp. 209–210. IEEE (2015). https://doi.org/10.1109/3DUI.2015.7131774