Computers & Education 139 (2019) 102–117

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Mobile virtual reality for musical genre learning in primary education

Edoardo Degli Innocenti[a], Michele Geronazzo[b,*], Diego Vescovi[a], Rolf Nordahl[b], Stefania Serafin[b], Luca Andrea Ludovico[c], Federico Avanzini[c]

[a] Dept. of Information Engineering, University of Padova, Padova, Italy
[b] Dept. of Architecture, Design, and Media Technology, Aalborg University, Copenhagen, Denmark
[c] Dept. of Computer Science, University of Milano, Milano, Italy

ARTICLE INFO

Keywords:
Mobile virtual reality
Music primary education
Music genre learning
Navigation
Spatial audio

ABSTRACT

Mobile virtual reality (VR) is increasingly becoming popular and accessible to everyone that holds a smartphone. In particular, digital didactics can take advantage of natural interaction and immersion in virtual environments, starting from primary education. This paper investigates the problem of enhancing music learning in primary education through the use of mobile VR. To this end, technical and methodological frameworks were developed, and were tested with two classes in the last year of a primary school (10 years old children). The classes were involved in an evaluation study on music genre identification and learning with a multi-platform mobile application called VR4EDU. Students were immersed in music performances of different genres (e.g., classical, country, jazz, and swing), navigating inside several musical rooms. The evaluation of the didactic protocol shows a statistically significant improvement in learning genre characterization (i.e., typical instruments and their spatial arrangements on stage) compared to traditional lessons with printed materials and passive listening. These results show that the use of mobile VR technologies in synergy with traditional teaching methodologies can improve the music learning experience in primary education, in terms of active listening, attention, and time. The inclusion of pupils with certified special needs strengthened our results.

1. Introduction

In many western countries, music education has undergone cuts of public funding, following a more general trend of deprioritization of music and other art-related activities (Dwyer). These cuts are unfortunate, given the many potential benefits that musical education may have for children. Music education includes both practical activities (e.g., learning how to play a musical instrument) and theoretical subjects (e.g., music theory and history). Previous research shows that music education is beneficial to improve spatio-temporal cognitive skills (Črnčec, Wilson, & Prior, 2006). Although there is some evidence that music education improves performance in other subjects such as mathematics, such benefits have not been demonstrated yet (Sala & Gobet, 2017).

Virtual Reality (VR) has been recently adopted to teach several subjects, including anatomy (Jang, Vitale, Jyung, & Black, 2017), architecture (Abdullah, Kassim, & Sanusi, 2017), art history (Casu, Spano, Sorrentino, & Scateni, 2015), neurosurgery (Pelargos et al., 2017), training and mentoring of educators (Lim, Han, Oh, & Jang, 2019), as well as scientific disciplines that require the use of a

* Corresponding author.
E-mail address: mge@create.aau.dk (M. Geronazzo).

https://doi.org/10.1016/j.compedu.2019.04.010
Received 6 July 2018; Received in revised form 19 April 2019; Accepted 20 April 2019
Available online 09 May 2019

laboratory (De Jong, Linn, & Zacharia, 2013).[1] Music education, conversely, has not been exposed yet to thorough research on the potential of VR approaches and applications, even if the sound and music computing community is starting to investigate the potential of VR for several applications (e.g., virtual and mixed reality instruments (Serafin, Erkut, Kojs, Nilsson, & Nordahl, 2016)).

In this paper we present a technical and methodological framework for music education in primary school based on a mobile VR environment. Specifically, we propose a novel VR application, named VR4EDU hereafter, which is used in this study to facilitate learning about musical genre. As an aspect of novelty with respect to other immersive environments in the musical field (see Section 3), here the focus is on music theory and history rather than music performance.

The paper is structured as follows: Sec. 2 defines the research questions, and Sec. 3 discusses the state of the art regarding research and applications in this field; in Sec. 4, technical details about the proposed framework are provided; while Sec. 5 describes the evaluation phase (participants, test protocol, statistics), Sec. 6 reports both quantitative and qualitative data; Sec. 7 addresses the key elements emerging from our experience, leading to final answers to research questions in Sec. 8, that also proposes a road map for future work.

2. Research questions

Starting from successful VR-based experiences in other educational fields discussed in the scientific literature, we aim to investigate the use of VR environments for music learning in primary school. It is worth noting that music education is a domain often refractory to technological innovation, being characterized by traditional one-to-one or frontal group lessons (Miendlarzewska & Trost, 2014). By looking through the lens of this specific domain, we aim at addressing some general research questions.

The first question (*RQ1*) is about the most effective way to integrate VR in an educational experience, in such a way to both maximize the learning outcomes and minimize potential disadvantages such as VR sickness and reduced possibilities of interaction in a peer-learning context.

The second research question (*RQ2*) is about young students' learning achievements, taking also in consideration children with special educational needs (SEN), whose performances could be affected by the adoption of a VR environment.

Finally, we want to investigate the user experience with technology-based learning tools, and specifically VR environments (*RQ3*). This question is specified in terms of students' perceived effort, engagement, motivation, and appreciation.

In our research, we partially address the Italian<>music national program for primary schools, with specific reference to a subset of learning goals, knowledge, skills, and activities.[2] Specifically, we deal with knowledge of musical instruments: VR4EDU is based on presenting and teaching eight different music genres to children. We address this learning goal by enhancing the listening experience, which is traditionally based on mono or stereo reproduction in the classroom: with VR4EDU the listening experience is individual, as music is delivered through earphones and augmented through immersive spatial audio. The main focus of the application is the learning of musical genres: VR4EDU exposes children to musical pieces that belong to different genres and can be listened and explored interactively. The corresponding learning activities exploit the use of movement with music: VR4EDU includes active body movements to let pupils move inside the 3D virtual environment and to encourage them to dance and move their bodies as they listen to the music.

Even though this research is focused on a very specific teaching subject, the information collected in the assessment phase may allow to infer some wider implications about the advantages of integrating emerging technologies in learning environments, about the impact of VR on students' performances, and about their experience with such tools. More specifically, we argue that our results may be generalized to those educational domains sharing similar characteristics: primary school students, theoretical subjects, and educational contexts that are still anchored to traditional teaching methodologies. Examples may range from art history to geography and to those STEM[3] subjects that can benefit from immersive learning environments.

3. Related work

The potential of new technologies to enhance student achievements and learning – if used appropriately – has been long recognized (Council, 2000). More specifically, the experimentation with VR technologies in primary, secondary, and higher education began in the early 1990s (Byrne & Furness, 1994; Youngblut, 1998), with the use of head-mounted displays, data gloves, and body suits. Overall previous research shows some encouraging results regarding the use of VR in educational settings (Dalgarno & Lee, 2010). A meta-analysis regarding the effectiveness of the use of VR in education (Merchant, Goetz, Cifuentes, Keeney-Kennicutt, & Davis, 2014) confirms its potential, even more when game-based environments are used as opposed to simulations. Moreover, the benefits are generally greater when students are tested immediately after the learning experience.

Researchers and educational practitioners have emphasized that the immersion provided by VR technologies offers strong benefits that can support education (Dede, 2009): it promotes learning through increased engagement of students, and facilitates the transfer of knowledge from the classroom to the real world; moreover, it fosters educational experiences that draw on "situated learning" and "learning-by-doing". The former definition refers to learning that takes place in a community of practice, in the same context in which it is applied (Lave & Wenger, 1991), while the latter postulates that active involvement of students in constructing new knowledge

[1] A valuable resource on VR and education is this site: https://vrschoolresearch.com/.

[2] A table summarizing the main program items is provided as supplementary material.

[3] STEM acronym identifies the following disciplines: science, technology, engineering and mathematics.

aids them in mastering, retaining, and generalizing such knowledge (Schank, Berman, & Macpherson, 1999). These approaches contrast with traditional instructional environments in which students learn by assimilation, i.e., by passively listening to an instructor.

Recent research (Christopoulos, Conrad, & Shukla, 2018), focused on the identification and taxonomy of the elements and the factors that affect learner engagement with virtual worlds, provides insights and guidelines that are particularly relevant to this work. Specifically, the authors emphasize the importance of having an "orientation" process aimed at familiarizing students with the virtual environment. Furthermore, they show that the coexistence of both the virtual and the traditional learning environment minimizes the drawbacks of each educational approaches.

3.1. Music learning in VR environments

Simulation has been applied, in various forms, as a tool for learning and training in many fields, such as architecture, medicine and sport, but still not fully explored in music. The point of view of musicians about simulation-based training has been discussed in (Aufegger, Perkins, Wasley, & Williamon, 2017), showing that musicians see benefits for developing, experimenting with and enhancing their performance skills within a simulated environment.

The increasing availability of affordable head-mounted displays has facilitated the development of several applications of VR for music. Most of such applications are targeted to playing VR musical instruments rather than learning about music theory and genre in VR.

As an example, the Music Room[4] is a collection of instruments. By using hand-held controllers, the platform enables interaction with a set of percussive instruments. The philosophy around the Music Room is that it acts as a MIDI controller for use in any Digital Audio Workstation (DAW). A similar application is Soundstage VR,[5] which, apart from interactive VR instruments, also includes a modular mix chain with a library of effects and processing, as well as a looping and recording stage for use in post-production or other media productions. The applications described above promote playing and interacting with VR musical instruments in order to produce music or to train rhythmical skills, and with a commercial purpose.

New musical interfaces based on VR have also been proposed in the academic world. In a recent work (Serafin, Adjorlu, Nilsson, Thomsen, & Nordahl, 2017), researchers report the state of the art and discuss the new perspectives of both virtual and augmented reality technologies in K-12 music education.

In (Hamilton & Platz, 2016), an abstract environment is presented where performers can collaboratively create music in VR. Berthaut and co-workers (Berthaut & Hachet, 2016) have also proposed several immersive interfaces for musical performances, arguing how the power of interactive 3D graphics, immersive displays, and spatial interfaces is still under-explored in domains where the main target is to enhance creativity and emotional experiences.

Over the years, several VR musical instruments have been designed with the main purpose of creating simulations or extensions of existing ones (Serafin et al., 2016). Moreover, virtual environments for learning to play traditional instruments have also been proposed in the past, e.g., flute (Johanna et al., 2010) and piano tutor systems (Chow, Feng, Amor, & Wünsche, 2013; Hwang, Son, & Kim, 2017), to name a few. In recent times, Orman et al. (Orman, Price, & Russell, 2017) investigated the feasibility of adopting VR learning environments to improve music conducting skills.

Augmented reality, where a traditional instrument is enhanced with technology such as visual, auditory and/or haptic feedback, appears to be more suitable than virtual reality for teaching to play musical instruments. This is mainly due to the fact that these are extremely sophisticated interfaces that have developed over years, while VR interfaces provide primarily an immersive visualization still having primitive controls in comparison with augmented instruments (Serafin et al., 2016). One can think at the importance of multisensory feedback and perception-action mechanisms involved in the perceived quality of musical instruments (Fontana, Papetti, Jrvelinen, & Avanzini, 2017), which is almost impossible to render in VR through currently available devices such as joysticks or datagloves.

Therefore several augmented reality based systems have appeared in the literature, such as Andante and Andantino (Xiao et al., 2014, 2016), where augmented reality is used to project animated characters on a piano keyboard. This not only engages children in playing the right note, but also on exploring the embodied characteristics of a music performance. Augmented reality has also been used to sensitize young children to abstract concepts of music, such as the musical notation or the idea of rhythm (Rusiñol, Chazalon, & Diaz-Chito, 2018).

3.2. Learning with sound in space

Spatial perception and cognition, especially related to navigation within an environment, rely on multimodal information. According to (Wiener, Bchner, & Hlscher, 2009), navigation within a real/virtual environment is based on three complementary mechanisms: (i) knowledge about a point in space such as landmarks or a destination, (ii) knowledge about a path to a destination ("route knowledge"), and (iii) integrated knowledge about the environment, i.e., cognitive-map like knowledge, or "survey knowledge".

In everyday life, people gain "on-line" cognitive mapping during exploration of an unknown environment, whether it is real or

[4] musicroomvr.com/.
[5] soundstagevr.com/.

virtual; spatial representation is thus updated in parallel while navigating (Burgess, 2006). Reviews of behavioral, neuroimaging, and electrophysiological studies suggest a close connection between goal-dependent and -independent representations of space (Chersi & Burgess, 2015). Place cells within the hippocampus provide a rapid associative memory connecting the goal and its context, i.e., the 3D environment, guiding navigation towards a desired destination. Moreover, the influence of context cues on a task is crucial also to reinforce memory (Smith & Vela, 2001). Multimodal spatial features can be learned in some extent without conscious, intentional focus resulting in an enhanced learning with no extra workload (Mandler, Seegmiller, & Day, 1977; Ragan, Bowman, & Huber, 2012).

Moreover, the influence of context cues on a task is crucial to reinforce memory in real-life scenarios (Smith & Vela, 2001). In line with this, immersive and interactive virtual reality technologies increasingly support the sense of personal, social and environmental presence. For the auditory channel in particular, the scientific literature often adopts the word “auralization” to refer to spatial audio rendering in virtual environments (Vorländer, 2007). Spatial audio technologies allow the direction of users attention and the enhancement of realism of the VR experience with positive influence on workload, performance, and presence score, as in the study of Bormann et al. (Bormann, 2005). In that work, the authors reported results of a search task in a VR environment where participants were more involved with the auditory aspects rather than the visual aspects when the object actively produced sounds.

The circular interaction between presence and emotions is well known in the scientific literature, leading to consider virtual reality as an affective medium (Riva et al., 2007), i.e. able to interact with affective states (Gorini & Riva, 2008) and memory processes (Sauzon et al., 2011). With regard to our study, a special connection with music must be drawn. Since music evokes strong emotions, it can also be involved in forming memories about information associated with a particular context, music genre or piece (Jäncke, 2008). The influence of music on our emotional and cognitive system can be modulated and enhanced by the valence rating (Eschrich, Mnte, & Altenmller, 2008). For all these reasons, the affective interactions between context-related memory and musical contents make active navigation in VR an affordable and effective tool for learning music.

4. The VR application

We developed an immersive VR listening experience enhanced with visually appealing 3D models of instruments. We designed two different VR Applications, the “Musical Labyrinth Exploration” and the “Room Learning Experience”. The former represents a training experience where pupils can learn how to move in the VR4EDU 3D environment, while the latter is the actual presentation of music genres as an enhancement to the school program.

This section presents the main elements and innovative aspects of the applications. All technical details related to the hardware-software implementation are available in the supplementary materials.

4.1. Financial feasibility

Most of the teaching applications developed so far make use of rather expensive VR devices. One of the reasons why VR technologies are beyond the reach of primary schools is financial feasibility (Moreno-Ger et al., 2010). However, this limitation is now being overcome thanks to the development of low-cost hardware technologies and the availability of simulation software packages. For the purposes of classroom activities, low-cost solutions such as cardboard-based VR headsets with smartphones can be made available to a classroom, facilitating interactive personalized immersive learning experiences (Amer & Peralez, 2014; Dede, 2009). Mobile VR technologies in particular can be easily introduced in classroom settings (Ball & Johnsen, 2016). In this paper we follow this approach, and use low-cost commercially available mobile VR headset in conjunction with free cross-platform software.

4.2. VR scenarios and interactions

The “Musical Labyrinth Exploration” (MLE hereafter) scenario is a 3D labyrinth that provides initial support to the navigation system, so that users have the opportunity to practice and get confident in exploring the 3D environment. To this purpose, the application proposes the following task: the user is asked to find four different colored areas in which four floating spheres are placed, representing sound sources of four different music tracks (see Fig. 1 for a schematic representation of the map).

This task is designed in order to encourage the user to explore the whole labyrinth, progressively learning to avoid walls and to walk in narrow and articulated paths. In order to help users understand how they are moving in the virtual environment, we implemented a chessboard-style (rather than uniformly-colored) floor. Moreover, in order to make music tracks easily distinguishable, we selected four different timbres: piano, classic guitar, violin, flute. To further encourage the exploration experience, tracks are made audible through walls so that the user can understand the direction of the incoming sound even when the floating sphere is not visible. Accordingly, users are instructed to pay attention to both the visual and the audio input for navigation, exploiting the sound source as a beacon sound in their way-finding task (Geronazzo, Bedin, Brayda, Campus, & Avanzini, 2016).

The “Room Learning Experience” (RLE herafter) scenario is composed by a limited rectangular space where the user can listen to a music piece representative of a specific musical genre. In addition, 3D representations of musical instruments of a specific genre are positioned in the space like they were placed on a performance stage, in the usual arrangement of a typical real performance. We decided to limit the available walking space in order to help the user focus on the instruments and their sound, minimizing distraction.

The user is presented by a welcome message introducing the room and an always-on message with the title of the musical genre of that room, along with its main characterizing instruments. This textual information can always be read at a glance, moreover the correct name of each instrument is also placed close to the corresponding 3D representation (see Fig. 2 for a screenshot).

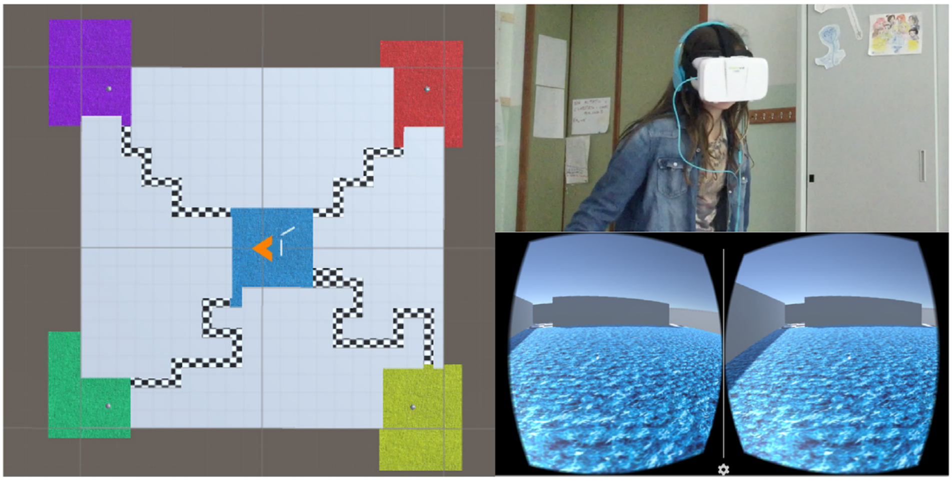


Fig. 1. Top view of the MLE labyrinth (left side) explored by a pupil (up-right side). The red, purple, green and yellow boxes are the four rooms and the light blue box is the starting center room. Connecting paths are depicted with chessboard squares. A small dot is visible in each room and represents the audio source. An orange arrow represents user position and orientation for which the first-person view is provided (down-right side). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

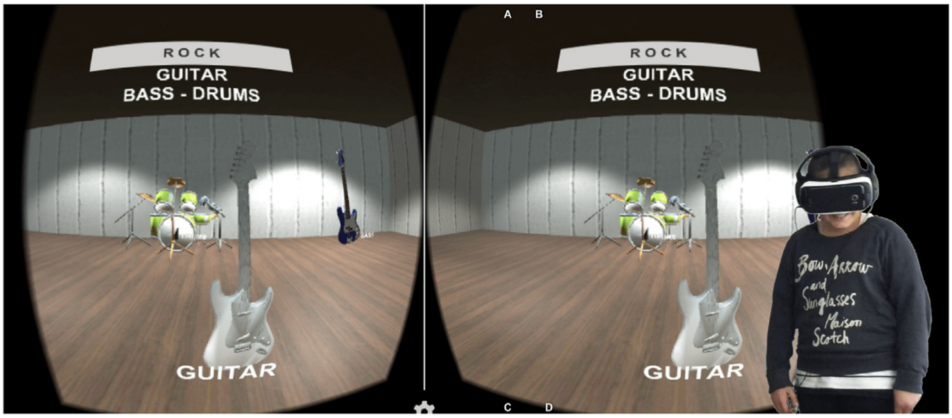


Fig. 2. Split screen game view from the RLE application. Here the main title is visible with the genre name and its main characterizing instruments.

The RLE application implements eight different "genre rooms", along with their respective characterizing instruments. The architectural space of each room does not vary among genres. Only title, labels and musical instruments (with their spatial arrangement) change.

4.3. Navigation

One of the key aspects of the VR4EDU application is the design of a reliable and intuitive navigation system able to provide spatial orientation cues in supporting usability and learning (Merchant et al., 2012). The adoption of a VR headset requires an effective "virtual locomotion" approach to be implemented, as the way people interact with the virtual environment is crucial for enhancing the learning experience and preventing problems like motion sickness and loss of balance.

The problem of virtual locomotion has been studied extensively (Steinicke, Visell, Campos, & Lécuyer, 2013). We considered several options, especially taking inspiration from common interfaces such as remote controllers and joysticks, as well as walking in place approaches (Usoh et al., 1999). Most of these solutions need a physical input interface, that would keep the users' hands constantly busy. In order to give the most natural and intuitive experience of movement for navigation and interaction in fixed world problem (Hürst & Helder, 2011), we designed a hands-free solution taking advantage of the inertial sensors of the VR headset and a set of simple body gestures. Specifically, we defined the following gestures:

- tilt head forward/backward (pitch rotation) → move ahead/back;
- tilt head right/left (yaw rotation) → slide to the right/left;
- rotate head around (roll rotation) → rotate view;

In order to make this solution useable and prevent instability, we implemented our virtual locomotion strategy in such a way that no tangential movement occurs if the tilting input is lower than a certain threshold. Moreover, we tuned the parameters in such a way that the user move forward at a maximum speed of about 25 m/s (see the accompanying supplementary materials). Our preliminary usability test with a group of five 10–11 years old children, who did not take part to the subsequent evaluation, showed that this maximum speed value grants a smooth navigation experience.

5. Evaluation

The effectiveness of the VR4EDU application in supporting music education was assessed through an evaluation with primary school children. The goal of the evaluation was to compare the learning outcomes of a class that used VR4EDU with those of a class that was exposed to traditional lessons only. The focus of the lessons was on the learning of musical genres.

5.1. Participants

Thirtysix children took part in the evaluation phase, in collaboration with their teachers. They all were 10–11 years old children attending the last year of primary school D. Alighieri in the Istituto Comprensivo Statale II "E.C. Davila" <> , Piove di Sacco (Padova, Italy) <> .

The first group (control group herafter) used traditional lessons only. The group was composed of 18 children (7 females and 11 males), four of them with SEN. They all took part to the whole evaluation.

The second group (VR group hereafter) used VR4EDU. The group was composed of 18 children (9 females and 9 males), six of them with SEN. They all took part to the whole evaluation.

5.2. Protocol

The experiment was conducted as a between-subject design: traditional lessons were conducted in both groups, VR group in addition used the VR4EDU application while control group allowed a comparison in learning performances. The didactic protocol consisted of four lessons for both groups, in a timespan of four weeks (one lesson per week). The main teacher followed both groups in all tests and lessons. Two technicians followed only VR group helping children to wear VR headsets and headphones.

Table 1 shows the musical genres considered in the evaluation, the corresponding musical pieces proposed to children, and the main associated musical instruments. The work-flow for the two groups is illustrated in Fig. 3.

During the first lesson, both groups carried out a pre-test (described in Sec. 5.3 below) to verify their previous knowledge. Students in VR group, in addition, became acquainted with VR4EDU and especially with its virtual locomotion approach: specifically, they explored the MLE application and were subsequently asked to draw the labyrinth that they discovered (some examples are reported in Fig. 4).

During the second lesson, both groups analyzed the following musical genres with the teacher: Blues, Classic Piano Trio, Country, Rock. Specifically, they analyzed their history, their characterizing instruments and their disposition on a stage. The lesson devoted about 15 min per genre. Students in the control group had to listen to a representative piece for each genre through a mono loudspeaker (about 2 min per genre), and to work on a form cutting out an outlined shape of instruments on paper, pasting and coloring. On the other hand, during the genre analysis in the VR group, students explored the room developed for each genre in the RLE application (about 2 min per genre), and after that they completed the corresponding form. It is worthwhile to notice that the duration of such exposure was short in order to limit cybersickness in the VR group, which has been shown to be a danger especially when using mobile VR devices (Moro, Štromberga, & Stirling, 2017). Cybersickness has been shown to have an impact on the learning experience and can present as nausea, disorientation, discomfort, headache fatigue, difficulty concentrating and problems with vision (Settgast, Pirker, Lontschar, Maggale, & Gütl, 2016).

The third lesson was similar to the second one: both groups followed the same work-flow but analyzed different music genres, namely Disco, Jazz, Folk, Swing.

Finally, during the fourth lesson both groups carried out a post-test (described in Sec. 5.3 below) to assess their learning performance and make a comparison between groups. At the end of the post-test, a questionnaire (described in Sec. 5.4 below) was

Table 1
Association between genre, song and typical musical instruments.

Genre	Piece	Main instruments
Blues	Sweet home Chicago	Cigar Box, Harmonica
Classic Piano Trio	Haydn Piano trio n. 39	Harpsichord, Violin, Viola
Country	Fly around my pretty little miss	Banjo, Violin
Disco	Togheter forever	Synth
Folk	Cincirinella teneva teneva	Tamburine, Accordion
Jazz	Take five	Saxophone, Contrabass
Rock	Satisfaction	El. Bass, El. Guitar, Drum
Swing	Singing on nothing	Trumpet

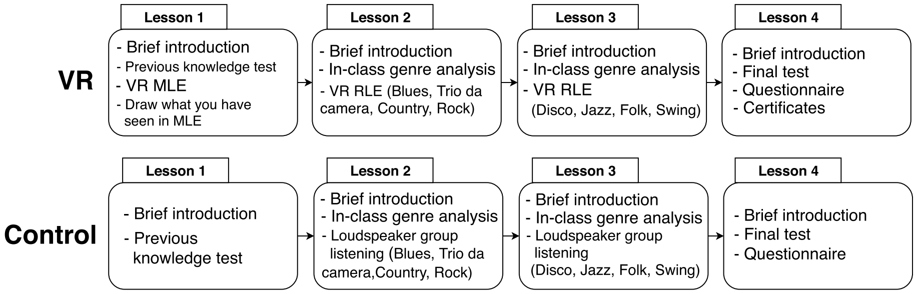


Fig. 3. Work-flow for the two groups.

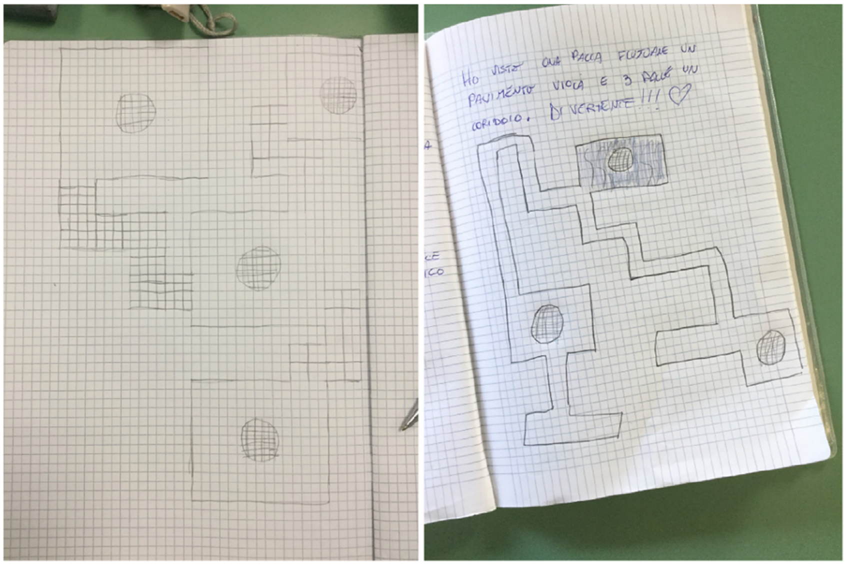


Fig. 4. Examples of drawings by children depicting the labyrinth of the MLE application (to be compared with Fig. 1).

proposed to both groups to collect qualitative data regarding the overall learning experience.

5.3. Quantitative evaluation and data analysis

The same test template was used both to determine previous knowledge in the first lesson (pre-test), and to assess the learning performance in the fourth lesson (post-test). The teacher always adopted such test in order to evaluate this part in the school program and consisted of three exercises:

1. **Genre identification**: the students listened to four pieces and, at the end of each, had to write down the identified genre;
2. **Association** between musical instruments and genre: the students had to link the images and names of seven musical instrument to the genre in which they were used;
3. **Spatial arrangement** of instruments in the performance space: the students had to indicate the position (close-up or background) of eight instruments arranged on three concert stages (e.g., location of electric guitar, electric bass and drum in Rock genre).

The pre-test and the post-test differed only in the choice of genres and instruments. This differentiation was introduced in order to not make the exercise trivial. Two genres (Rock and Swing) were present in both tests, while the remaining two were different.

In the pre-test, Exercise 1 asked to recognize the four genres Rock, Jazz, Classic Piano Trio, Swing; Exercise 2 asked to link the seven instruments Harmonica, Banjo, Tamburine, Trumpet, Synth, Accordion, Electric Guitar; Exercise 3 asked to arrange on stage, writing names in the intended position, the eight instruments Electric Guitar (front) Drums and Bass (back) for Rock, Violin and Viola (front) Harpsichord (back) for Classic Piano Trio, Saxophone and Contrabass (both front) for Duo Jazz. One example of question for each Exercise is shown in Fig. 5. [6]

On the other hand, in the post-test Exercise 1 asked to recognize the four genres Rock, Swing, Disco, Blues; Exercise 2 asked to link

[6] Full tests are provided in the supplementary materials.

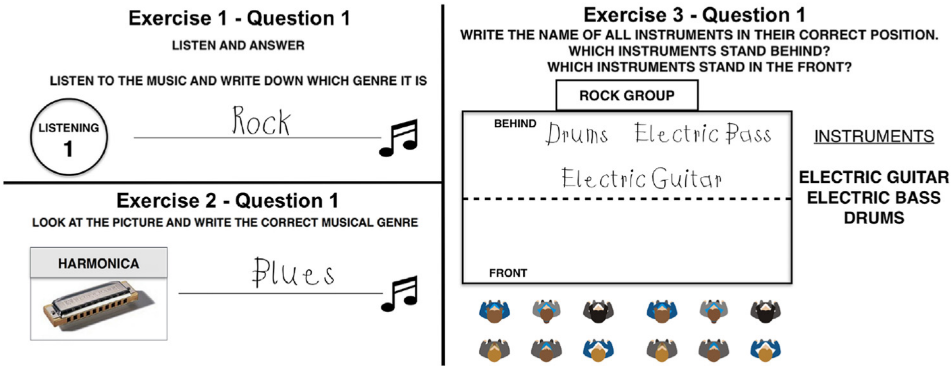


Fig. 5. Examples of question and answer for each exercise. Total number of responses for each exercise: 4 (Exercise 1), 7 (Exercise 2), and 8 (Exercise 3). In particular, each question of Exercise 3 requires a number of responses equals to that of the expected instruments on stage.

the seven instruments Harmonica, Saxophone, Tamburine, Trumpet, Contrabass, Accordion, Harpsichord; Exercise 3 asked to arrange on stage the same eight instruments of the pre-test.

Answers to the three Exercises were equally weighted in the final score of the test, leading to a total of 19 responses (4, 7, and 8 answers for Exercises 1, 2, and 3, respectively). Accordingly, the final score was the percentage of correct answers. In addition, the percentage of correct answers in each single exercise was also computed in order to assess peculiarities among different learning aspects. Statistical evaluation was performed following a mixed-model design: a one-way ANOVA on a between-group factor with two levels of didactic method (i.e., traditional and VR-supported lessons), and a within-group factor with two levels of test (i.e., pre- and post-test), was performed on test scores. A preliminary analysis on score distributions were subjected to Levene's test for homoscedasticity with no violation for the homogeneity of variances assumption, and inspections of linear model residuals for score values showed that normality assumption was not violated according to a Shapiro-Wilk test. Post-hoc analysis on interactions/contrast with Bonferroni correction procedures on p-values provided pairwise statistical comparisons in test scores between groups and class tests.

5.4. Qualitative evaluation

The didactic experience was also subjectively evaluated with an *ad-hoc* questionnaire which is summarized in Table 2. Seven questions were given to both groups, while two additional questions were given to VR group only. Questions q1-4 and q8-9 used a 5-point Likert scale (1–"Not at all", 2–"A little", 3–"Moderately", 4–"Quite a bit", 5–"Very much"), with increasing check-box size in order to be easily understandable to a child (Wrzesien & Alcaiz Raya, 2010).

Statistical analysis of questionnaire items was performed by comparing response data distributions between control and VR-supported groups with a Monte Carlo estimation of the Fisher's exact p-value for count data (with 10^4 replicates); moreover, Cronbach's alpha were computed based on participants' scores in order to establish whether each items is reliably measured.

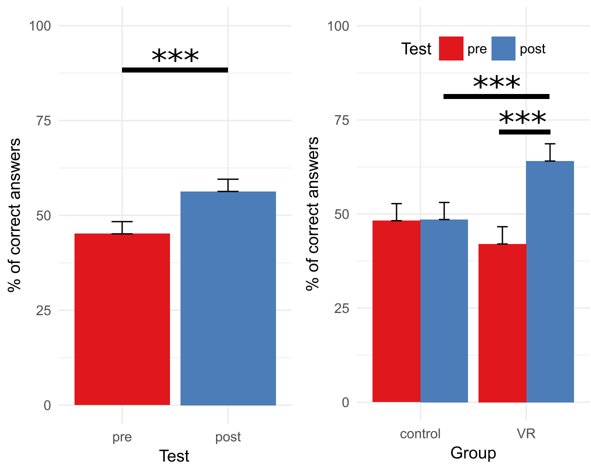
6. Results

6.1. Quantitative data

A one-way ANOVA on final scores revealed a statistically significant interaction between didactic method and type of test [$F(1,31) = 26.66$, $p \ll .001$, $\eta^2 = .20$]. An overall score increment for all pupils after the didactic experience was also statistically significant according to pairwise comparison ($p \ll .001$, see Fig. 6. a for trends) of pre- (M:45% SD:8%) and post- (M:56% SD:13%)

Table 2
Questionnaire's questions and scales.

	Question	Scale	
q1	How hard did you find to learn the different music genres?	5-point Likert	
q2	How much were you engaged during the activities?	–	
q3	Would you like to discover more musical genres?	–	
q4	How much did you enjoy the music lessons on musical genres?	–	
q5	What did you like? Write the aspects that you enjoyed during these lessons on musical genres.	open-ended	
q6	What did you dislike? Write the aspects that you did not like, or what would you change in these music lessons.	–	
q7	Do you know virtual reality? What do you think?	–	
q8	How much did you enjoy doing lessons with virtual reality?	5-point Likert	VR only
q9	How much is virtual reality useful?	–	–



(a) Type of test (b) Didactic method

Fig. 6. Global statistics (average + standard error) for percentage of correct answers in the evaluation tests grouped by (a) pre- and post-test, and (b) didactic methods. Asterisks and bars indicate, where present, a significant difference (*: $p < .05$, **: $p < .01$, ***: $p < .001$ at *post-hoc* test).

performances. However, looking at the contrast in didactic methods, only VR group exhibited a statistically significant improvement ($p \ll .001$, see Fig. 6. b, pre - M:41% SD:8%, post - M:64% SD:11%), while control group did not ($p = .99$, pre - M:48% SD:7%, post - M:48% SD:10%). Moreover, pairwise comparisons on contrast in test type revealed no statistical significance in the pre test ($p = .99$, control - M:48% SD:7%, VR - M:42% SD:8%), denoting an equal starting level in the two groups; a statistically significant difference between groups in the post-test ($p \ll .001$, see again Fig. 6. b, control - M:48% SD:9%, VR - M:64% SD:11%) suggested a differentiation in learning due to the two didactic actions.

Tables 3 and 4 summarize the same statistical analysis performed for each single exercise, in order to assess different aspects of the learning process.

For Exercise 1, a one-way ANOVA on percentage of correct answers revealed a statistically significant interaction between didactic method and type of test. It is worthwhile to notice that performance in Exercise 1 was on average lower in the post- (M:25% SD:25%) than in the pre-test (M:40% SD:23%) for the control group; however, looking at the contrast in didactic methods, no statistically significant improvements were detected (see Fig. 7a) for both groups (VR, pre - M:30% SD:14%, post - M:44% SD:25%); the difference in the mean could be induced by the differentiation in the 4 genres between pre- and post-tests. Pairwise comparisons on contrast in test type denote a comparable starting knowledge in the two groups (pre, control - M:40% SD:23%, VR - M:30% SD:14%); on the other hand, a statistically significant improvement in the post-test (control - M:25% SD:25%, VR - M:45% SD:25%)) suggests a differentiation in learning due to the didactic actions.

For Exercise 2, the ANOVA on percentage of correct answers revealed statistical significance in the type of test only. In particular, the contrast within didactic methods was statistically significant for VR group (see Fig. 7b) denoting an improvement in learning from pre- (M:32% SD:12%) to post- (M:54% SD:23%) with the support of VR4EDU.

Finally, the ANOVA on percentage of correct answers of Exercise 3 reported a statistically significant interaction between didactic method and type of test. Contrast in didactic methods exhibited a statistically significant improvement between pre- (M:57% SD:15%) and post- (M:87% SD:13%) tests only in the VR group (see Fig. 7c). Moreover, pairwise comparisons on contrast in test type denotes an equal starting level in the two groups similarly to previous exercises (control - M:68% SD:14%, VR - M:57% SD:15%), and a statistically significant improvement in the post-test (control - M:67% SD:16%, VR - M:83% SD:13%) suggests a differentiation in

Table 3
Results of ANOVA statistical tests performed on each single exercise. Asterisks indicate, where present, a significant difference (*: $p < .05$, **: $p < .01$, ***: $p < .001$).

Source		ANOVA			
		MS	F (1,31)	p	η^2
Exercise 1	Factor 1: didactic method	.031	.479	.49	.01
	Factor 2: type of test	.000	.001	.94	.00
	interaction	.341	9.94	**	.10
Exercise 2	Factor 1: didactic method	.085	2.68	.11	.04
	Factor 2: type of test	.433	15.8	***	.18
	interaction	.062	2.25	.14	.03
Exercise 3	Factor 1: didactic method	.011	.560	.46	.01
	Factor 2: type of test	.259	11.9	**	.13
	interaction	.290	13.4	***	.16

Table 4
Results of pair-wise comparisons performed on each single exercise. Asterisks indicate, where present, a significant difference (*: $p < .05$, **: $p < .01$, ***: $p < .001$ at *post-hoc* test).

Group	Dependent variable	Paired Differences			t value	df	p
		Test	Estimated contrast	SE			
control	Exercise 1 - score	pre-post	-.147	.064	−2.31	31	.06
	Exercise 2 - score	–	.101	.057	1.77	31	.17
	Exercise 3 - score	–	-.007	.050	-.146	31	.99
VR	Exercise 1 - score	pre-post	.141	.066	2.15	31	.08
	Exercise 2 - score	–	.223	.059	3.80	31	**
	Exercise 3 - score	–	.258	.052	4.96	31	***
Pre	Exercise 1 - score	control-VR	.010	.077	1.29	56	.41
	Exercise 2 - score	–	-.010	.060	-.176	61	.99
	Exercise 3 - score	–	.106	.051	2.09	61	.08
Post	Exercise 1 - score	control-VR	-.188	.077	−2.41	56	*
	Exercise 2 - score	–	-.133	.060	−2.22	61	.06
	Exercise 3 - score	–	-.159	.051	−3.14	61	**

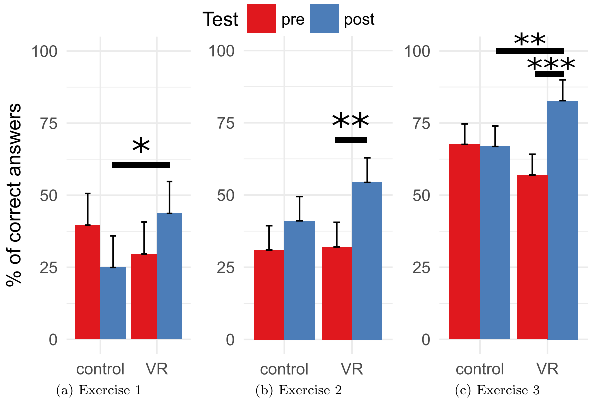


Fig. 7. Global statistics (average + standard error) for percentage of correct answers grouped by didactic methods in pre- and post-tests, for (a) Exercise 1, (b) Exercise 2, and (c) Exercise 3. Asterisks and bars indicate, where present, a significant difference (*: $p < .05$, **: $p < .01$, ***: $p < .001$ at *post-hoc* test).

didactic actions.

6.2. Children with special educational needs

According to the school's teaching body, our study includes also children with SEN: 4 in the control group, and 6 in the VR group. Since previous global results do not consider the possible effect for those children in the experimental measures, we perform a further

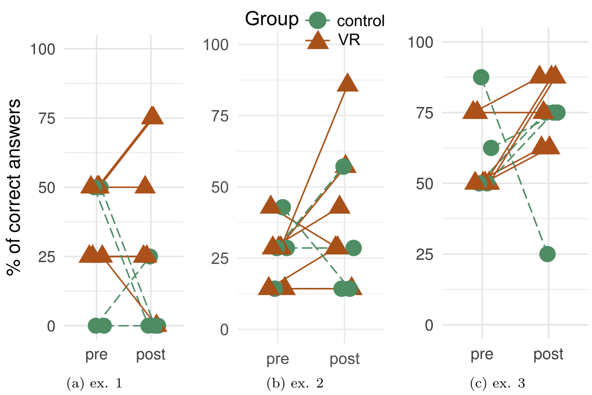


Fig. 8. Percentage of correct answers for each child with SEN, grouped by didactic methods in pre- and post-tests, for (a) Exercise 1 (b) Exercise 2, and (c) Exercise 3.

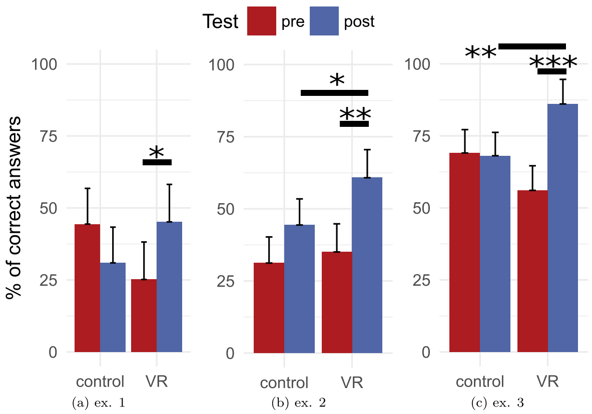


Fig. 9. Global statistics (average + standard error) for percentage of correct answers without SEN children, grouped by didactic methods in pre- and post-tests, for (a) Exercise 1, (b) Exercise 2, and (c) Exercise 3. Asterisks and bars indicate, where present, a significant difference (*: $p < .05$, **: $p < .01$, ***: $p < .001$ at *post-hoc* test).

analysis considering performances of those children and their classmates, separately.

Fig. 8 depicts individual scores of SEN children for each exercise, visually divided by group. Performances between pre- and post-tests denote differences between groups. Pupils in the VR group increased or maintained their overall score in all exercises, with the exception of two pupils in Exercises 1 and 2 with a decrease <25% in score, while one or more pupils in the control group worsened their score with a decrease of >25% in each exercise. Moreover, average standard deviations (deltas) between pre- and post-performances for each exercise are in line with descriptive global statistics, showing no improvements for the control group and a positive effect in the VR group. Specifically, for control group the average deltas were − 18% (± 38%), 0% (± 23%), and 0% (± 42%) in Exercises 1, 2, and 3, respectively. On the other hand, for the VR group the average deltas were 4% (± 19%), 17% (± 25%), and 18% (± 15%) in Exercises 1, 2, and 3, respectively.

A one-way ANOVA on the final scores without considering children with SEN revealed a statistically significance interaction between didactic method and type of test [$F(1,21) = 27.58$, $p \ll .001$, $\eta^2 = .27$]. An overall score increment for all pupils was also statistically significant according to pairwise comparison ($p \ll .001$, pre - M:46% SD:9%, post - M:60% SD:11%). However, looking at the contrast in didactic methods, only the VR group exhibited a statistically significant improvement ($p \ll .001$, pre - M:42% SD:10%, post - M:68% SD:6%), while the control group did not ($p = .99$, pre - M:50% SD:6%, post - M:52% SD:9%). Moreover, pairwise comparisons on contrast in test type revealed statistical significance in both the pre- ($p < .05$, control - M:50% SD:6%, VR - M:42% SD:10%) and post- ($p \ll .001$, control - M:52% SD:9%, VR - M:68% SD:6%) tests, suggesting a differentiation in both starting knowledge and learning effect. Considering that the control group started from a higher level of previous knowledge and experienced no improvements, the results of the VR group are even more significant.

The statistics on each exercise without SEN children resulted in increased consistency with the global statistics (see Fig. 9). The sole exception was the missing statistical significance in Exercise 1 between the post-tests of the two groups ($p = .07$, control - M:31% SD:25%, VR - M:45% SD:23%).

6.3. Qualitative data

We analyzed the qualitative data obtained by subjective questionnaires. The internal consistency reliabilities using the Cronbach's α coefficient was 0.82 for all items with $\alpha = .84$ and $\alpha = .74$ for the control and VR-supported groups, respectively. The Fisher's Exact Test for count data with simulated p-value resulted in no statistical differences among the two groups of children for the first three questions (q1: $p = .50$, q2: $p = .38$, q3: $p = .28$). On the other hand, q4 provided a statistically significant difference ($p \ll .001$).

The general impression was that VR group was more involved in the learning process than control group. In the following, we discuss q1-4 group-wise; for the sake of readability, the response distributions are reported in square brackets, respecting scale order (see also Fig. 10 for a graphical representation of data distribution and density using likert R-package [7]):

q1: lessons on musical genres were slightly lighter for VR group, [Control: (1, 6, 8, 1, 1); VR: (2, 9, 5, 0, 0)];

q2: in accordance with q1, students' engagement in VR group was slightly higher, arguably thanks to the novelty of the VR4EDU framework, which elicited curiosity among children [Control: (0, 1, 9, 2, 5), VR: (0, 1, 5, 6, 4)];

q3: both groups gave positive answers, indicating a shared curiosity and interest in discovering other musical genres [Control: (1, 2, 2, 2, 10), VR: (0, 2, 0, 6, 8)];

q4: VR group enjoyed lessons much more [Control: (1, 0, 3, 8, 5); VR: (0, 0, 0, 1, 15)].

Interestingly, answers to q2 for the control group are close to a bimodal distribution with centers in 3 and 5, indicating a marked distinction between a large group (%̃50) of neutrally engaged pupils and a smaller one that was highly engaged in the lessons. On the

[7] https://cran.r-project.org/package-likert.

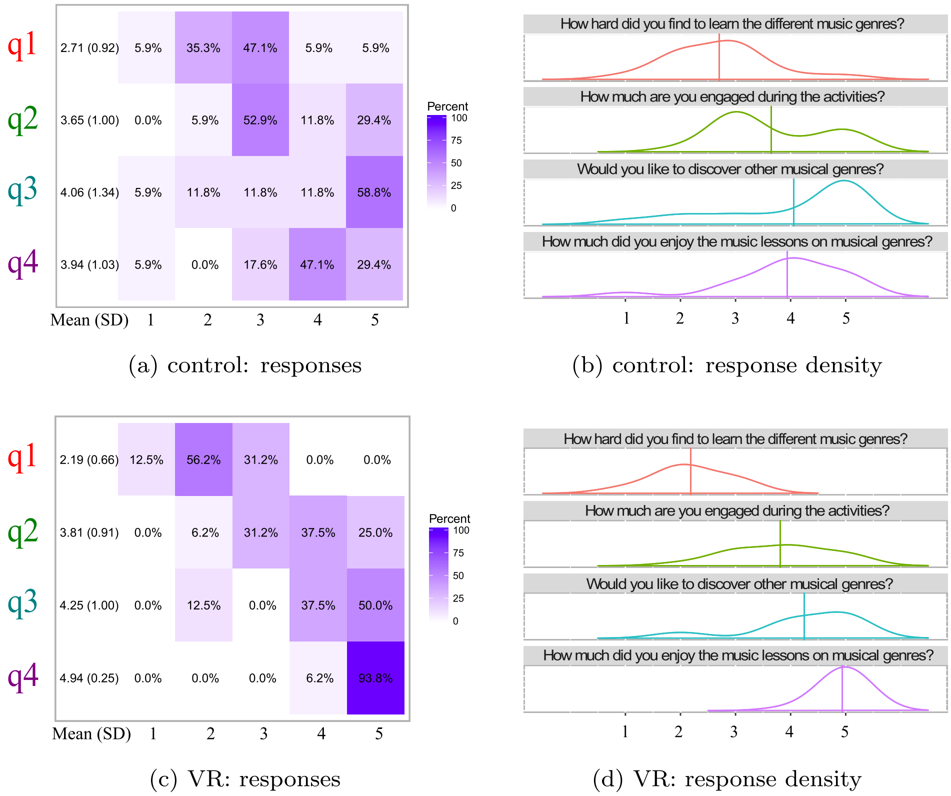


Fig. 10. Response distributions for Likert questions q1-4, grouped by didactic method.

other hand, students' engagement was distributed smoothly in the VR group.

The open-ended questions q5-7 supported trends and results of the previous ones. In particular for the VR group, answers to q5 denoted many positive aspects: level of enjoyability of the experience, and rendering details from which they were most impressed during navigation (e.g. visual textures, musical pieces in VR, etc.). Students in the VR group had fun exploring the virtual environments. Activities in the traditional lessons, i.e., listening, reading, coloring, cutting and pasting the form while the teacher explained genres, were perceived as a one single enjoyable experience and the VR exploration like a smooth extension of such experience. Analyzing negative aspects in q6, only one student in the VR group said that he did not like the virtual locomotion approach, while others focused on music genres that they did not like; most students simply stated that they liked everything and nothing needed to be changed. Regarding q7, only five students in the VR group reported that they were already familiar with VR. Those pupils who explicitly expressed their idea on VR reported high expectations of this technology for music teaching, teaching methodologies in general, and playful experiences.

In the control group, answers to q5 were similar to those of VR group regarding learning of new musical genres and instruments. While analyzing q6, we noticed negative comments about activities in the traditional lesson (e.g., writing, cutting and pasting) that were not stressed in the VR group. Finally, we did not notice any differences in answers to q7 compared to VR group, with the sole exception of a smaller number of comments connecting VR with music lessons. This is in accordance with the fact that children in control group did not use VR4EDU and were not aware of experiments with the VR group.

The last two questions, q8-9, were asked to the VR group only. Both the questions had extremely positive responses (q8: (0, 0, 0, 1, 15); q9: (0, 0, 0, 3, 13)).

7. General discussion

The results of our evaluation suggest that VR-supported lessons allow a more effective learning experience with statistically significant improvements compared to pre-existing knowledge and mid-term learning of traditional lessons alone (Fig. 6). It is worthwhile to notice that the VR group included a larger number of children with SEN; if these students are omitted from the analysis, the pre-test shows that the VR group had lower previous knowledge of the topic compared to the control group, especially related to Exercises 1 and 3 (see also average values in Figs. 7 and 9).

Although this work is not primarily focused on special educational needs, the presence of SEN children in both groups allowed to analyze the results separately for the two populations. This analysis provided some interesting indications and showed that the proposed approach can ensure an equally effective experience for the whole class, thus supporting the goals of inclusive education

(Ainscow & Sandill, 2010). No firm conclusions can be drawn about this aspect, since the current study involved a statistically small number of SEN children, for which individual differences were not assessed. Moreover, different proportions of SEN children in the classroom may also affect the results. For these reasons, further investigations should take into account these factors as relevant variables.

A first key element that explains the global differentiation between control and VR groups (further corroborated by extremely positive answers to q8 and q9) is the well-known evidence in scientific literature of the benefits in learning due to affordances of spatial and affective interactions in virtual reality (Dalgarno & Lee, 2010). The immersion and navigation in the VR scenario supported children attention, and the proposed context elicited a link among previously acquired knowledge with traditional methods, spatial information and emotional memories. On the other hand, pupils may have benefited from a novelty effect which is difficult to measure on such a short (few minutes) experience with VR. Our work provides a first step towards the integration of short VR learning experiences in the classroom: the long-term effectiveness should be assessed using more educational materials and across a larger numbers of lessons. Investigating long-term learning is a challenging task which would require collection of new data and the development of new VR experiences, which should still be tightly connected to the traditional education program.

A specific characterization of the didactic action can be drawn by looking at each exercise. Exercise 1 was the most difficult to perform, to such an extent that control group performed worse in the post-test than in the pre-test. On the other hand, VR group increased its performance between pre- and post-. These results should not surprise, considering that pupils listened to one single piece of music for each genre and for a limited amount of time, not allowing any robust classification in their mind. Exercise 2 revealed a statistically significant contribution of VR-supported lessons in improving VR group performances with respect to previous knowledge, and a tendency to provide better learning compared to control group. These results suggest that VR experience strengthened the connection between instruments and genre. According to Sing et al. (Singh et al., 2015), anchoring to spatial locations can support understanding and retrieving of contents and associations, and improving of memory. Multimodal environmental features of a place, i.e. stage, can aid storing related information. Thus, the simple recalling of a place in memory can help remember all such information (Tulving & Osler, 1968).

Performances in Exercise 3 were generally higher than in previous exercises. Learning improvements were reported only for the VR group, with statistically significant differences in both type of test and didactic methods. These results reinforce the above discussion on Exercise 2, supporting the importance of spatial learning and immersion.

The above results are even more significant when considering that the benefits of VR4EDU were obtained with just 2 min of VR experience per musical genre. Considering that 4 musical genres were presented in one hour class, this means that pupils used VR4EDU for just 8 out of 60 min (13% of the time).

Unpacking the effects of the two factors discussed above, i.e. spatial learning and immersion, is not easy. In particular, since listening attentively to music is greatly aided by the quality of the audio playback, one may argue that the differences in audio quality between the two groups (immersive stereo over headphones versus mono loudspeaker) influenced the results to a greater extent than spatial learning: in this respect, a fairer choice would be for the control group to experience high-quality audio through headphones. On the other hand, this would still be a passive listening experience which does not necessarily lead to immersion, due to the lack of a direct connection to children movements and interactions (Maes, Leman, Palmer, & Wanderley, Wanderley). In order to further investigate this aspect, an improved experimental protocol should allow the control group to experience different audio technologies for auralization (e.g. stereo, 5.1, Ambisonics, etc. (Vorländer, 2007)), that provide a variety of auditory information and different levels of interactivity and immersion.

Similarly, the spatial design proposed in this study only scratches the potential of spatial learning. More complex designs may be conceived, in which other spatial aspects can be learned about musical genres. As an example, virtual rooms might contain peculiar elements for each genre (e.g. specific furniture, architectural elements, posters, etc.) in order to foster associative memory. On the other hand, these additional virtual objects could also become elements of distraction and divert child's attention from listening to spatialized music. Pupils have shorter attention spans compared with adults (Cowan, Fristoe, Elliott, Brunner, & Saults, 2006), and one should carefully consider any distracting elements that could have a negative impact on attention. For these reasons, at this stage of our research we chose to minimize the number of independent variables and to use a small set of visual spatial aspects (e.g. the relative location of the different instruments in a performance). We also limited the virtually walking area to a medium size room in order to support attention and focus on few aspects.

Results from the questionnaire show that negative comments in q6 were higher in number for the control group, being weakly related to the topic of music genres, but only to the traditional didactic method. Note however that traditional lessons were conducted in both groups: students in VR group perceived the lessons as more interesting because they were more involved, being better prepared to learn and face workload. Also, the questionnaire did not highlight any negative effects of the VR experience, such as VR sickness or lack of communication. We argue that during this short and rich experience the strong demand for attention fully absorbed the pupils' attention, and subsequently fostered them to share their impressions with peers during the remaining portion of lesson time. In this respect, mechanisms of informal peer learning that are typical of shared work may have been even reinforced by the VR experience. However these considerations will need to be corroborated through the collection of new data while investigating long-term learning, as discussed above.

This study also provides some technical insights on the use of VR in primary education, which can inform developers and instructional designers. We obtained the following simple empirical guidelines for creating simple virtual rooms.

- The interface should be minimal and easy to comprehend, with particular attention to text font with high contrast improving readability. For VR4EDU application we used Arial font as suggested by Woods et al. (Woods, Davis, & Scharff, 2005).

- Always-on information on screen should be used with care in order not to overload visual information (Ruijten, Kruyt-Beursken, & IJsselsteijn, 2018), which would cause distraction from auditory information; this issue is particularly relevant for children who have never used a VR headset before. As an example, we adopted the Cardboard reticle pointer[8] that enlarges the pointing circle if it overlaps a musical instrument. Other enhancements (e.g, highlighting edges of focused instruments, increasing dimensions, or triggering a special animation) have to be evaluated on a case-by-case basis.
- Light control is also an important issue; in RLE, we adopted spotlights over instruments in order to focus children attention (Fig. 11a) similarly to the work of Khan and colleagues aimed at directing the visual attention on large wall-sized displays (Khan, Matejka, Fitzmaurice, & Kurtenbach, 2005); in MLE, we chose a natural light that resemble a sunny day in order to evoke a positive feeling of playing outside (Fig. 11b).

8. Conclusions

In this work we presented the VR4EDU framework, a VR-based learning tool that exploits spatial learning and immersive listening to support music lessons in primary school. Even though experimentation in the field of music didactics is currently at an early stage, and most of the existing applications are targeted to music performance rather than theoretical aspects, our work demonstrated the applicability of VR also to the latter. In particular, to our knowledge VR4EDU is the first VR application targeted at musical genre recognition and learning. Some of the results of this work allow to infer some wider implications for VR and education. In particular, our assessment of the user experience through the final questionnaire is largely independent on the specific domain of music education and music genres (see questions from q5 on).

Let us recall the research questions listed in Section 2. The potential drawbacks involved in the adoption of VR in an educational environment – including both physical problems such as VR sickness and pedagogical issues such as the isolation of pupils during VR experiences – have been addressed by proposing mixed educational activities, where reasonably short VR sessions accompany traditional lessons. This educational approach proves to be effective, as shown by the qualitative and quantitative assessment presented in Section 5 (*RQ1*). We obtained statistically significant learning improvements with VR-supported music lessons. Our results show that the teaching of music can benefit from this kind of enhancement of traditional didactic methods. Moreover, pupils with special needs were able to achieve the same or even better results compared to classmates, overcoming personal, writing/reading, or attention difficulties (*RQ2*). Finally, the user experience with a VR environment coupled with traditional teaching activities was positively rated by students in terms of perceived effort, engagement, motivation, and appreciation (*RQ3*).

Future developments will be focused on improving the didactic protocols and the VR framework, according to the discussion provided in Sec. 7, and in the limits of available resources. In particular, once the ratio between the number of available VR devices and pupils will be 1:1, the protocol will be updated accordingly in order to maximize the time spent in VR without interruption. On the technical side, we believe that increasing the level of realism and details of the musical instruments and stages with dynamic 3D models of performers should be considered and evaluated in terms of immersion and location awareness. We are also currently improving the audio experience, with more complex multi-track recordings and different musical layers that will be rendered in three dimensions, exploiting novel 3D audio rendering algorithms that optimize the individual listening experience (Geronazzo, Peruch, Prandoni, & Avanzini, 2017; Romigh & Simpson, 2014).

The proposed framework and guidelines can also be extended to other topics in the music school program, and to younger children (age <10) in order to develop and validate a more prolonged exposure to VR-supported lessons. As an example, a similar topic is learning the connection between instruments and their family (e.g., the guitar is a chordophone, drums are membranophones, etc.). The proposed protocol can be adapted accordingly, focusing on classification by sound production mechanisms, rather than genres. As a further example, the same approach can be applied to the learning of the relationship between sounds and their environments, i.e. learning how sounds such as a school bell, a washing machine, traffic jam, are related to specific environments like a school, a house, an urban street. Also in this case, increased immersion through the use of personalized 3D spatial sound rendering can create a strong connection, thus resulting in a reinforcement for learning.

Ad-hoc evaluation with children with special needs should also be taken into account in order to explore the use of VR technologies as compensatory tools for teachers, supporting learning and inclusion. We also plan to expand the testing of the developed technologies to other cultures and curricula beyond the one investigated in this paper.

Acknowledgements

This study was supported by the internationalization grant of the 2016-2021 strategic program "Knowledge for the World" awarded by Aalborg University to Michele Geronazzo.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compedu.2019.04.010.

[8] developers.google.com/vr/unity/refereznce/class/gvr-reticle-pointer.

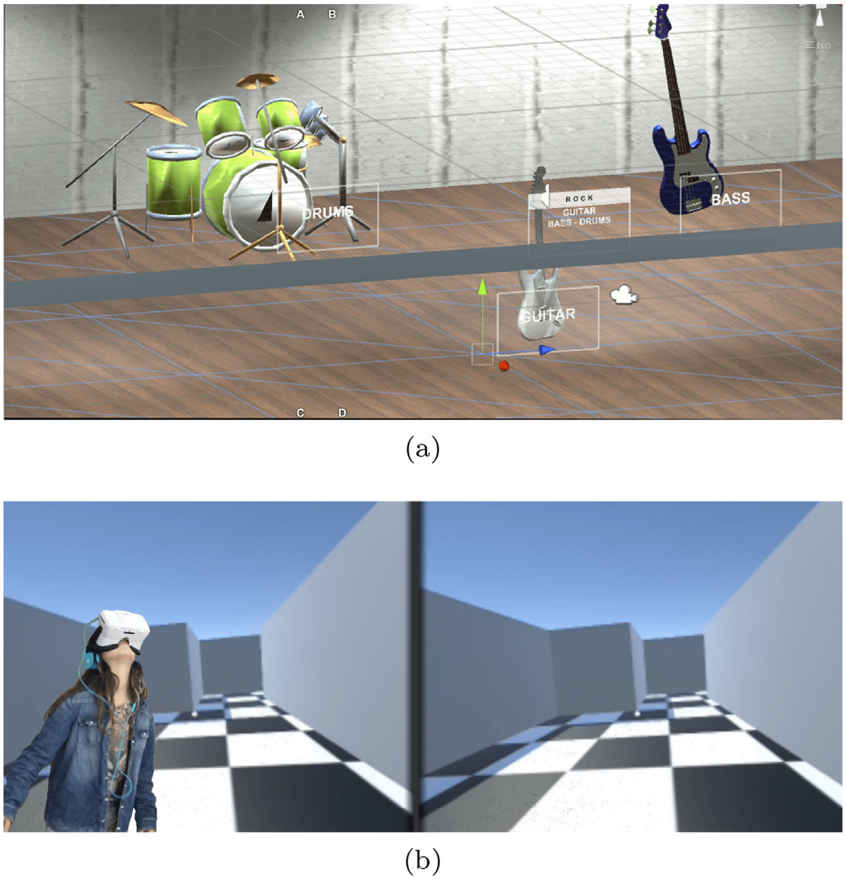


Fig. 11. (a) Scene view from Unity project of VR4EDU application of the rock genre room. Instruments names labels are visible in front of them. (b) Split screen game view from Unity project of VR4EDU application of musical labyrinth exploration.

References

Abdullah, F., Kassim, M. H. B., & Sanusi, A. N. Z. (2017). Go virtual: Exploring augmented reality application in representation of steel architectural construction for the enhancement of architecture education. *Advanced Science Letters, 23*(2), 804–808.
Ainscow, M., & Sandill, A. (2010). Developing inclusive education systems: The role of organisational cultures and leadership. *Int. J. of Inclusive Education, 14*(4), 401–416.
Amer, A., & Peralez, P. (2014). Affordable altered perspectives: Making augmented and virtual reality technology accessible. *IEEE global humanitarian technology conference (GHTC 2014)* (pp. 603–608). .
Aufegger, L., Perkins, R., Wasley, D., & Williamon, A. (2017). Musicians perceptions and experiences of using simulation training to develop performance skills. *Psychology of Music, 45*(3), 417–431.
Ball, C., & Johnsen, K. (2016). An accessible platform for everyday educational virtual reality. *Everyday virtual reality (WEVR), 2016 IEEE 2nd workshop on, IEEE* (pp. 26–31). .
Berthaut, F., & Hachet, M. (2016). Spatial interfaces and interactive 3d environments for immersive musical performances. *IEEE Computer Graphics and Applications, 36*(5), 82–87.
Bormann, K. (2005). Presence and the utility of audio spatialization. *Presence, 14*(3), 278–297.
Burgess, N. (2006). Spatial memory: How egocentric and allocentric combine. *Trends in Cognitive Sciences, 10*(12), 551–557.
Byrne, C., & Furness, T. A. (1994). Virtual reality and education. *Exploring a new partnership: Children* (pp. 181–189). Teachers and Technology.
Casu, A., Spano, L. D., Sorrentino, F., & Scateni, R. (2015). Riftart: Bringing masterpieces in the classroom through immersive virtual reality. *Eurographics Italian chapter conference* (pp. 77–84). .
Chersi, F., & Burgess, N. (2015). The cognitive architecture of spatial navigation: Hippocampal and striatal contributions. *Neuron, 88*(1), 64–77.
Chow, J., Feng, H., Amor, R., & Wünsche, B. C. (2013). Music education using augmented reality with a head mounted display. *Proceedings of the fourteenth Australasian user interface conference-volume 139* (pp. 73–79). Australian Computer Society, Inc.
Christopoulos, A., Conrad, M., & Shukla, M. (2018). Increasing student engagement through virtual interactions: How? *Virtual Reality,* 1–17.
Council, N. R. (2000). *How people learn: Brain, mind, experience, and school* (Expanded edition). National Academy Press.
Cowan, N., Fristoe, N. M., Elliott, E. M., Brunner, R. P., & Saults, J. S. (2006). Scope of attention, control of attention, and intelligence in children and adults. *Memory & Cognition, 34*(8), 1754–1768.
Črnčec, R., Wilson, S. J., & Prior, M. (2006). The cognitive and academic benefits of music to children: Facts and fiction. *Educational Psychology, 26*(4), 579–594.
Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology, 41*(1), 10–32.
De Jong, T., Linn, M. C., & Zacharia, Z. C. (2013). Physical and virtual laboratories in science and engineering education. *Science, 340*(6130), 305–308.
Dede, C. (2009). Immersive interfaces for engagement and learning. *Science, 323*(5910), 66–69.
M. C. Dwyer, Reinvesting in arts education: Winning America's future through creative schools, President's Committee on the Arts and the Humanities.
Eschrich, S., Mnte, T. F., & Altenmller, E. O. (2008). Unforgettable film music: The role of emotion in episodic long-term memory for music. *BMC Neuroscience, 9*, 48.
Fontana, F., Papetti, S., Jrvelinen, H., & Avanzini, F. (2017). Detection of keyboard vibrations and effects on perceived piano quality. *Journal of the Acoustical Society of*

America, 142(5), 2953–2967.
Geronazzo, M., Bedin, A., Brayda, L., Campus, C., & Avanzini, F. (2016). Interactive spatial sonification for non-visual exploration of virtual maps. *International Journal of Human-Computer Studies, 85*, 4–15.
Geronazzo, M., Peruch, E., Prandoni, F., & Avanzini, F. (2017). Improving elevation perception with a tool for image-guided head-related transfer function selection. *Proc. Of the 20th int. Conference on digital audio effects (DAFx-17), Edinburgh, UK* (pp. 397–404). .
Gorini, A., & Riva, G. (2008). Virtual reality in anxiety disorders: The past and the future. *Expert Review of Neurotherapeutics, 8*(2), 215–233.
Hamilton, R., & Platz, C. (2016). Gesture-based collaborative virtual reality performance in carillon. *Proceedings of the international computer music conference 2016* (pp. 337–340). .
Hürst, W., & Helder, M. (2011). Mobile 3d graphics and virtual reality interaction. *Proceedings of the 8th international conference on advances in computer entertainment technology* (pp. 28). New York, NY, USA: ACE '11ACM 1–28:8.
Hwang, I., Son, H., & Kim, J. R. (2017). Airpiano: Enhancing music playing experience in virtual reality with mid-air haptic feedback. *World haptics conference (WHC), 2017 IEEE* (pp. 213–218). IEEE.
Jäncke, L. (2008). Music, memory and emotion. *Journal of Biology, 7*(6), 21.
Jang, S., Vitale, J. M., Jyung, R. W., & Black, J. B. (2017). Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. *Computers & Education, 106*, 150–165.
Johanna, K., Romero, G., Andres, D., Lopez, R., Luengas, L. A., Carlos, J., et al. (2010). Virtual flute: Electronic device that uses virtual reality to teach how to play a flute. *IEEE education engineering (EDUCON)* (pp. 211–216). .
Khan, A., Matejka, J., Fitzmaurice, G., & Kurtenbach, G. (2005). Spotlight: Directing users' attention on large displays. *Proceedings of the SIGCHI conference on human factors in computing systems, CHI '05, ACM, New York, NY, USA* (pp. 791–798). .
Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge university press.
Lim, D. H., Han, S. J., Oh, J., & Jang, C. S. (2019). Application of virtual and augmented reality for training and mentoring of higher education instructors. *Handbook of research on virtual training and mentoring of online instructors* (pp. 325–344). IGI Global.
P.-J. Maes, M. Leman, C. Palmer, M. M. Wanderley, Action-based effects on music perception, Frontiers in Psychology 4.
Mandler, J. M., Seegmiller, D., & Day, J. (1977). On the coding of spatial information. *Memory & Cognition, 5*(1), 10–16.
Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in k-12 and higher education: A meta-analysis. *Computers & Education, 70*, 29–40.
Merchant, Z., Goetz, E. T., Keeney-Kennicutt, W., Kwok, O.-m., Cifuentes, L., & Davis, T. J. (2012). The learner characteristics, features of desktop 3d virtual reality environments, and college chemistry instruction: A structural equation modeling analysis. *Computers & Education, 59*(2), 551–568.
Miendlarzewska, E. A., & Trost, W. J. (2014). How musical training affects cognitive development: Rhythm, reward and other modulating variables. *Frontiers in Neuroscience, 7*.
Moreno-Ger, P., Torrente, J., Bustamante, J., Fernández-Galaz, C., Fernández-Manjón, B., & Comas-Rengifo, M. D. (2010). Application of a low-cost web-based simulation to improve students practical skills in medical education. *International Journal of Medical Informatics, 79*(6), 459–467.
Moro, C., Štromberga, Z., & Stirling, A. (2017). Virtualisation devices for student learning: Comparison between desktop-based (Oculus Rift) and mobile-based (Gear VR) virtual reality in medical and health science education. *Australasian Journal of Educational Technology, 33*(6), 1–10.
Orman, E. K., Price, H. E., & Russell, C. R. (2017). Feasibility of using an augmented immersive virtual reality learning environment to enhance music conducting skills. *Journal of Music Teacher Education, 27*(1), 24–35.
Pelargos, P. E., Nagasawa, D. T., Lagman, C., Tenn, S., Demos, J. V., Lee, S. J., et al. (2017). Utilizing virtual and augmented reality for educational and clinical enhancements in neurosurgery. *Journal of Clinical Neuroscience, 35*, 1–4.
Ragan, E. D., Bowman, D. A., & Huber, K. J. (2012). Supporting cognitive processing with spatial information presentations in virtual environments. *Virtual Reality, 16*(4), 301–314.
Riva, G., Mantovani, F., Capideville, C. S., Preziosa, A., Morganti, F., Villani, D., et al. (2007). Affective interactions using virtual reality: The link between presence and emotions. *CyberPsychology and Behavior, 10*(1), 45–56.
Romigh, G. D., & Simpson, B. D. (2014). Do you hear where I hear?: Isolating the individualized sound localization cues. *Frontiers in Neuroscience, 8*, 370.
Ruijten, P. A. M., Kruyt-Beursken, E., & IJsselsteijn, W. A. (2018). Towards the simplicity of complex interfaces: Applying ephemeral adaptation to enhance user performance and satisfaction. In J. Ham, A. Spagnolli, B. Blankertz, L. Gamberini, & G. Jacucci (Eds.). *Symbiotic interaction* (pp. 86–97). Cham: Springer International Publishing.
Rusiñol, M., Chazalon, J., & Diaz-Chito, K. (2018). Augmented songbook: An augmented reality educational application for raising music awareness. *Multimedia Tools and Applications, 77*(11), 13773–13798.
Sala, G., & Gobet, F. (2017). When the music's over. Does music skill transfer to children's and young adolescents' cognitive and academic skills? A meta-analysis. *Educational Research Review, 20*, 55–67.
Sauzon, H., Arvind Pala, P., Larrue, F., Wallet, G., Djos, M., Zheng, X., et al. (2011). The use of virtual reality for episodic memory assessment: Effects of active navigation. *Experimental Psychology, 59*(2), 99–108.
Schank, R. C., Berman, T. R., & Macpherson, K. A. (1999). Learning by doing, Instructional-design theories and models. *A New Paradigm Of Instructional Theory, 2*, 161–181.
Serafin, S., Adjorlu, A., Nilsson, N., Thomsen, L., & Nordahl, R. (2017). Considerations on the use of virtual and augmented reality technologies in music education. *K-12 embodied learning through virtual & augmented reality (KELVAR), 2017 IEEE virtual reality workshop on, IEEE* (pp. 1–4). .
Serafin, S., Erkut, C., Kojs, J., Nilsson, N. C., & Nordahl, R. (2016). Virtual reality musical instruments: State of the art, design principles, and future directions. *Computer Music Journal, 40*(3), 22–40.
Settgast, V., Pirker, J., Lontschar, S., Maggale, S., & Gütl, C. (2016). Evaluating experiences in different virtual reality setups. *International conference on entertainment computing* (pp. 115–125). Springer.
Singh, G., Bowman, D. A., Hicks, D., Cline, D., Ogle, J. T., Johnson, A., et al. (2015). Designing a mobile augmented reality system for scaffolding historical inquiry learning. *2015 IEEE international symposium on mixed and augmented reality - media, art, social science, humanities and design* (pp. 9–14). .
Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review, 8*(2), 203–220.
Steinicke, F., Visell, Y., Campos, J., & Lécuyer, A. (2013). *Human walking in virtual environments.* Springer.
Tulving, E., & Osler, S. (1968). Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology, 77*(4), 593–601.
Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., et al. (1999). Walking¿ walking-in-place¿ flying, in virtual environments. *Proceedings of the 26th annual conference on Computer graphics and interactive techniques* (pp. 359–364). ACM Press/Addison-Wesley Publishing Co.
Vorländer, M. (2007). *Auralization: Fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality* (1st ed.). Incorporated: Springer Publishing Company.
Wiener, J. M., Bchner, S. J., & Hlscher, C. (2009). Taxonomy of human wayfinding tasks: A knowledge-based approach. *Spatial Cognition and Computation, 9*(2), 152–165.
Woods, R., Davis, K., & Scharff, L. (2005). Effects of typeface and font size on legibility for children. *Am. J. of Psychological Research, 1*(1), 86–102.
Wrzesien, M., & Alcaiz Raya, M. (2010). Learning in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the E-Junior project. *Computers & Education, 55*(1), 178–187 00154.
Xiao, X., Puentes, P., Ackermann, E., & Ishii, H. (2016). Andantino: Teaching children piano with projected animated characters. *Proceedings of the the 15th international conference on interaction design and children* (pp. 37–45). ACM.
Xiao, X., Tome, B., & Ishii, H. (2014). Andante: Walking figures on the piano keyboard to visualize musical motion. *The int. Conf. On new interfaces for musical expression (NIME)* (pp. 629–632). .
Youngblut, C. (1998). Educational uses of virtual reality technology. *Tech. rep*. Alexandria, Va: Institute for Defense Analyses.