Immersive and participatory serious games for heritage education, applied to the cultural heritage of South Tyrol.

Alessandro Luigini¹, Monica Parricchi¹, Alessandro Basso¹, Demis Basso²,³

¹ Free University of Bozen, Faculty of Education, Viale Ratisbona, 16, 39042 Bressanone-Brixen (BZ), Italy.  
² Free University of Bozen, Cognitive and Educational Laboratory (CESLab) Viale Ratisbona, 16, 39042 Bressanone-Brixen (BZ), Italy.  
³ Centro de Investigación en Neuropsicología y Neurociencias Cognitivas (CINPSI Neurocog), Universidad Católica del Maule, Talca, Chile.

Abstract. Heritage education is an activity that is increasingly present in the educational curricula of schools and museums. Cognitive mechanisms and Representation devices must be thoroughly analysed and designed in order to promote the creation of didactic paths to foster effective learning experiences. Immersive visualization technologies are well suited for gamification applications and the technological and economic accessibility of VR HMD (head-mounted display) viewers makes these technologies particularly attractive for the development of potentially more widespread methodologies. This article will describe an educational path, and the relative experimentation, on the cultural heritage focused on the production of the typical bread of the Val Pusteria area - and the rural life around it. The project was aimed at primary school children and was based on a serious game in Virtual Immersive Reality.

Keywords: Virtual Reality, heritage education, digital heritage, serious game.

1 Introduction

Over the past decade, the use of digital technologies in education has increased significantly, due to the recent development of devices and falling prices [1][2][3][4]. In particular, the use of serious games seems to be increasingly widespread in the context of Digital Heritage [5][6][7][8][9][10][11][12].

Serious games are not a recent innovation, and the most consolidated and authoritative definition is that of Abt [13], according to which games with deliberate educational intent can be considered serious games. These are games that players engage in for purposes other than entertainment. There have been many other interpretations [14][15][16][17]. However, such positions may appear as limited or often doubtful [18][19]. More generally, it may be considered that the most important distinction between entertainment games and serious games is the presence, in the latter, of educational content. [20].

A wide review of the literature indicates that serious games follow a typically constructivist approach, with the learner at the centre of the narrative [21]. According
to Osterman [22], constructivist pedagogical strategies should be based on the following concepts: 1. involve the learner; 2. provide opportunities to explore, articulate and represent knowledge; 3. challenge existing conceptual views and raise awareness of problems; and 4. allow learners to test the effectiveness of new points of view. The objective of an educational path in a formal context is not to entertain but to provide a tangible experience of knowledge. Serious games have the merit of involving students and motivating them to invest energy in learning tasks [19]. They promote extrinsic and intrinsic motivation [23].

The field of digital heritage seems to be one of the preferred areas of development for the use of serious three-dimensional and immersive games, as is professional training, medical training and simulations. [14][6][24]. One of the reasons is undoubtedly the need to visualize architectural or urban environments, in the case of educational paths that have to do with Tangible Cultural Heritage (TCH), while in the case of Intangible Cultural Heritage (ICH), the content is represented in many different ways, not necessarily visual or three-dimensional. It is well known that games offer a suitable environment for situated learning because they can take place in virtual environments that are reminiscent of or represent the real context in which they occur [25]. This aspect is fundamental in ICH education, where the teaching and learning process should be rooted in the context in which ICH is traditionally carried out, and the context itself plays an important role and has a value [26].

The serious games mentioned in the scientific literature use a three-dimensional environment in about half of the cases, and three-dimensional environments are more common in applications for students aged 12 and over [9][6][27]. A great deal of data is therefore available to assess the effectiveness of 3D-based serious games for secondary school students (age 12+) but almost none for primary school students (ages 6-11). There may be some scepticism towards the use of head-mounted displays (HMD) for students aged 6 to 11 years, considering this age group too young for the use of stereoscopic devices. Moreover, the diffusion of HMD in schools remains limited, while on-screen devices (desktop or mobile) are available in all schools. Such issues could be applied to immersive technologies too – such as scenarios and 360° videos – for which there is little or no experimental data available for these age groups (Rupp et al. 2016).

This work focuses on the use of immersive virtual reality technologies useful for the development of an educational path organized in the form of a serious game, dedicated to pupils of primary schools II and IV (7 and 10 years). The theme of the serious game presented here is teaching the cultural heritage, and the case study deals with the traditional bread of the Puster Valley and the rural tradition. The ICH included the peasant folk tradition that revolves around the traditional bread of the Puster Valley. Moreover, immersive digital environments were chosen because they offered greater possibilities for personalization and for adaptation to the context of the game, and because learning in a realistic environment could lead the educational path to success, if participants could experience a sense of familiarity.

Several didactic paths were designed, developed and experimented, all based on a User Centered Design (UCD) methodology. However, this article will focus on one didactic path in particular. The chosen proposal was dedicated to primary school pupils, using immersive technologies on a 3D basis and in the form of serious games. The age of the sample was chosen to fill the gap in the literature, while collecting evidence about
the possible use of immersive technologies in schools. This also influenced the choice of device: Oculus Rift (Facebook Technologies Ireland Limited) was selected because of its relatively low price, which could make this technology accessible in educational contexts. Moreover, particular attention was paid to adapt the features of the serious game to school children, for example by limiting their exposure to immersive vision and designing a game mechanism that does not risk making the student feel isolated. These aspects do not seem to have been considered in previous literature about serious games, and therefore deserve investigation.

2. The contribution of virtual reality on heritage education

The cultural tradition has associated the transmission and acquisition of knowledge with the symbolic-reconstructive way given by the study while the perceptive-motor model has always been considered of secondary importance [28]. New technologies tend to shift the communication and learning of knowledge from the first way to the second, more immediate, allowing all those involved to act directly and interactively on the objects of knowledge. The evolution of educational technologies has constantly led pedagogical and didactic research to rethink learning models according to new visions of the subject, the classroom and the school [29].

Technologies are not substituting, but integrative in teaching: it is necessary to consider digital as a re-mediation of reality and not a factor of discontinuity, that is, a reconfiguration in another key of the elements of everyday reality [30].

This centrality of the body does not originate from a direct physical relationship with the "natural" real, but from the comparison with man-made artifacts that constitute an anthropocentric habitat based on human cultural resources [31]. The feeling of presence, in the sense of being immersed in the environment we are observing, given by the activities in virtual reality is what defines it as an "extension of humans" [32].

Technologies related to virtual reality are progressively presenting increasing potentials in the field of education and training. These technologies are linked to human cognitive functioning, to the acquisition, processing and communication of knowledge by implementing learning processes. The drive to reformulate learning paradigms comes from the invention of technologies that allow an extraordinary diffusion of the symbolic medium and its support [33]. New universes of experience and knowledge are thus delineated, since the products of the most recent technologies are tools elaborated by human evolution in which the symbolic-technological, rational and representative dimension has become capable of producing articulated and complex constructs.

Regarding teaching at school, virtual reality constitutes an enhancement of the experience, providing experiential information, also acting as an "accessory" of the book or of direct experience and acting as an instrument of interacting discovery in digital environments [34]. This methodology allows to create a new potential for classroom use, as this project seeks to explore, as it is able to increase the real world of educational content and to create new and motivating ways for students with the opportunity to interact and compare with the surrounding environment [35].
Some aspects of contemporary learning theories can be traced back to this educational evolution [36] and they have inspired the design of educational path that we present:

- **Constructivist learning** – by using Virtual Reality in a way that encourages students to engage on a deeper level with the tasks, concepts and resources being studied through the use of information overlays, students can make deep and lasting connections within their knowledge base.

- **Situated learning** – authentic and contextualised learning is enabled by embedding educational experiences within the real-world environment and by bringing the real world into the classroom.

- **Games-based learning** – Virtual reality systems can be used to facilitate immersive games-based learning by creating a digital narrative, placing students in a role, providing authentic resources and embedding contextually relevant information.

- **Enquiry-based learning** – by offering a means to electronically gather data for future analysis and provide virtual models situated within a real-world context. Virtual Reality supports enquiry by providing information that is contextually relevant to the topic being investigated.

The combination of the basic knowledge of an object, such as a tree or monument, with the environment in which the object is located makes learning processes engaging, stimulating and collaborative.

In an integrated sense, technology allows educators to create a scenario, provide specific information, introduce characters into the learning experience and incorporate data. Virtual reality allows the user to see the real world, with virtual objects superimposed or composed with the real world, enhancing the physical and visual environments of an individual.

The virtual reality in the teaching process, in particular concerning heritage education, allows to:
- realize scenarios beyond the theoretical description;
- combine theoretical information with practical-experimental activities also through a more playful method;
- “learning by doing”, without real consequences in case of errors;
- use tags and labels (hotspots) to create links, including visual ones, that are easier to share and understand;
- modeling objects in various scenarios;
- realize and/or use projects and paths of museum/archaeological fruition with immersive experience.

A key pedagogical affordance of Virtual Reality, crucial for study on heritage education [27], is the ability to replicate virtual objects and environments - in our case the South Tyrolean traditional farms - , allowing students to better understand the properties and relationships of objects or environment that would be difficult to be effectively examined in the everyday life [37]. Although other technologies may perform the same function, re-dimensioning in systems provides the learner with a clear representation of spatial and temporal concepts and the additional advantage of contextualizing the relationship between the virtual object and the real-world environment.
3. Methods

3.1. Participants

The sample included 15 students in second grade (7-8-year-old) and 21 students in fourth grade (9-10-year-old) at the primary school “Rosmini” located in Brixen, in the Province of Bozen, South Tyrol, Italy. They were all examined, before they participated, to exclude any serious diseases. They all had normal or corrected-to-normal vision. Informed consent for their participation was obtained from their parents and from the Principal of their School.

3.2. Materials

3.2.1 The serious game design

The implementation of the serious game required in-depth reflection and a search for innovative solutions from a procedural point of view, but accessible in both technological and economic terms.

Given that the target group were children in the 2nd and 4th grades of primary school (respectively about 7-8 years old and 9-10 years old), the game had to consider their personal cognitive structure, time of adaptation and attitude towards collaboration. In order to adapt the device to the children, several aspects congruent to the theory of cognitive load were considered. In particular, the redundancy of textual information was avoided in favour of visual information; the entire serious game was set up with a sequencing process. The "tasks" required of the child were broken down, then re-aggregated into smaller groups (chunking) and presented in modular units with the same internal sequential structure. This sequential structure was organized according to a narrative principle [38]. In particular, the seven components that build a story – Event-tokens, Narrator, Narrative appetite, Past time, Structure, Agency, Purpose and Reader – seem to highlight important factors to be taken into account in the design of an educational path. In the design of our scenario, the concept of narrative appetite – [38] p. 541: the ability of a narrative to feed the desire to discover what will happen next and how the story will end – played a central role.

Several observational points were organized for each environment: the first series dedicated to the observation of the environment itself, without further stimuli, and only the last - following the activation of the "play" button – to answering questions (fig 1). This sequencing of the internal structure - and its deconstruction - made it possible for the children-players to maintain their concentration when necessary and to be free to explore the visual space, when possible.

The Serious Game was developed in seven successive scenarios, which lead children to visit a farm from the outside and from the wheat fields, through the kitchen, the bedroom, the living room (Stube), the barn and the roof, before returning outside to the outdoor oven. Each scenario is played by a different observer: for this reason, the group is composed of 7 students.
Fig. 1. This diagram illustrates a game session: you can see all the scenarios and characteristics of each one of them; in case of error or difficulty on the part of the operator (identified by ID12), the observers intervene avoiding moments of impasse and uncertainty by the main player.

Each environment has at least two points of view: the first, dedicated to observation, and the second, dedicated to the gameplay. The reasons for such a distinction lie in the need to reduce the potential distraction generated by the desire to explore during the game: in digital environments - immersive as well as non-immersive - each user is naturally led to explore the space (by motor or visual means) to understand its characteristics and to build relationships between their body and the environment. Thus, the first point of view - dedicated only to visual exploration - becomes the moment when the user enters into contact with the digital environment and, for subsequent thresholds of detail, pays attention to increasingly detailed information. In this way, probably, the activity to be carried out in the second point of view - the game itself -
can be more focused and more consistently "represent" the knowledge and skills of the player. Once a player has reached the point of view containing the game, he must activate it with a hidden button – located under the feet – that is visible only in the case of the passage of the pointer. Attention was necessary to avoid unintentional activations, which were frequent in beta testing.

When the game was activated, the players had to demonstrate knowledge and skills they acquired during their visit to the Folksmuseum in Teodone, by responding to quizzes or by completing predetermined activities (such correctly selecting the ingredients needed to bake rye bread). The transition to the next environment depended on providing a correct solution to the quiz, setting up the dichotomous structure stage by Norris et al. [38, p. 542].

In every environment, the players were able to visually explore rural life through the HMD (fig. 2), in analogy with their experience during the visit to the Museum of Teodone: environments, materials, furniture, furnishings and landscape were reconstructed through a philological study of the heritage of South Tyrolean traditional farms. The conclusion of the game was an external scenario, in which the teacher of the class summarized some of the content of the path, acting as a "pedagogical condenser" of the path itself: this expedient was fundamental to close the construction of the narrative path, to satisfy the narrative appetite enunciated by Norris et al. [38, p. 541].

Fig. 2. An HMD Oculus Rift was used for the experimentation, but the educational paths can also be edited for other devices, both HMD and mobile or desktop.

Particular attention was required in developing the visual system, both to maintain a balance of the cognitive load, and to limit or eliminate the effects of kinetosis or cybersickness – for further details: [39][40][41] –. An increasing number of cases has
been observed of people who, during or after the use of digital stereoscopic viewers, exhibit symptoms such as nausea, dizziness, headaches and increased sweating.

The psychological and physiological sciences have been studying the phenomenon for a long time [42][43][44]. This is because the onset of these symptoms has been observed not only in digital stereoscopic vision, but in situations such as traveling by car, train or airplane, or in astronauts during space flight.

Especially for children aged 2 to 12 years, who seem to be the subjects who most easily present these symptoms [42], attention to the smallest detail is important to reduce or eliminate the manifestation of these symptoms.

The design responses that we have elaborated will reduce or eliminate the risks listed above through certain specific settings of the representation device. First, we have chosen to limit the interaction of the player to the visual system, developing the serious game on the basis of 360° static images - rendered on the basis of digital models - and using a teleportation locomotion system similar to the one used, for example, in the Google Street view system, which may already be known and therefore "familiar" to the recipients of the project. Walking-based systems would probably have required greater body involvement by the children, and in the case of locomotion managed by touchpad controllers, an increased risk of discrepancy between visual and motor stimuli.

To move from one point to another, to activate phases in the game and interact with the quizzes, preference was given to a "point and click" system that can be activated both with a pointer sensitive to the movements of the VR viewer and with the touchpad controller (but only for 4th grade students). This choice has allowed users to reduce adaptation times to an extremely low threshold and to enter the game in an almost natural way.

3.2.2 Workflow for the serious game scenarios design

In order to achieve all the goals that we set for ourselves to characterize the digital environments, it was necessary to design and build three-dimensional environments in their entirety.

The initial goals were: a high level of realism, in order to promote the adaptation of the students to the VE; control of the elements featured in the scenes, in order to avoid the distractions possible with images of real environments which contain too many objects that are of little interest to the game activity. For the second goal we proceeded with a semantic and visual distinction between the game elements and the ambience elements, without going so far as to make the environment unnatural. For the first goal we followed a very elaborate workflow that allowed a considerable level of realism.
Fig. 3. Graphic workflow. Development phases and correlation with the use of the software involved in the project.

The development pipeline includes a series of procedures that bring this methodology closer to that used in the editing of next-generation video games, where the graphic quality and fluidity of use are essential to making the 3D resources workable on interactive platforms, making the gaming experience more natural and immersive. The digital modelling of the architectural environments required the simultaneous use of two distinct approaches: the first focused on the metric-figurative survey of a farmhouse of the same type as the one visited by the students; the second focused on the realization of 3d assets (furniture and objects) based on a philological study of the tools and furnishings to be found in a farmhouse, taken from historical documents.

The case study then developed two synchronic paths for the digital survey:

- Photo-modelling from photographic data sets acquired in real sample environments (since we did not have access to the Museum of Teodone, we chose farms of the same type, but accessible elsewhere)

- Manual modelling by interpretation based on historical, graphic and photographic documentation of the interiors and the various furnishings.

The first phase of the workflow consisted of photo-modelling using Agisoft Metashape and Meshroom software: the result was the point cloud and the subsequent mesh. At an operational level, the workflow we adopted was based on the semantic deconstruction of three-dimensional elements using certain subdivision criteria, such as material similarity or other taxonomic features. This procedure allowed interoperability between software environments and the transition to software such as Rhinoceros 3d and MOI 3d1, for the architectural elements, and to Maxon Cinema 4D, for the objects contained in the scenarios. The subsequent UV mapping2 was carried

---

1 These programmes are modeling software known as NURBS, an acronym that stands for Non-Uniform Rational Basis-Splines, which allow the production of high precision 3d models. This editing methodology is based on extrusion and envelope operations of curves and iso-curves, therefore two-dimensional elements drawn in vector mode, from which 3d models are generated.

2 UV mapping consists of matching a flattened mesh with an image. Each vertex of the three-dimensional object will thus have a set of two-dimensional coordinates shared with the image, which can then be associated with its faces, to become visible in 3D space. The UV coordinates therefore act as a bridge between the two-dimensional space of the images (textures) and the three-dimensional space of the mesh.
out with Pixologic ZBrush, a digital sculpture software often used for cinema graphics, which also allowed an optimal automatic retopology of the model. This process, which made use of auto-generative triangulation algorithms, was necessary for the transition from NURBS modelling software to polygonal modelling software. The process also allowed a correct attribution of the normal polygonal and, using the management functions of the Voxel algorithms, allowed the optimization of 3D resources, texturing and related creation of a UV Map corresponding to the new mesh generated.

An automatic retopology procedure was therefore necessary. This is generally a very complex phase that can be completed with good results relying on various automatic methods within the Pixologic software, using the functions related to the management of Voxel algorithms and polygonal division that have allowed a balanced optimization of 3D resources, texturing and the related creation of a UV Map corresponding to the new mesh generated, distinguishing the structural and solid parts from the different decorative parts and merging the resources associated with materials sharing the same characteristics. Cinema4D and ZBrush, linked by means of a plug-in bridge (GoZ), were used to generate various texture channels that simulate the detail for each surface element.

---

3 The topology in computer graphics is to be considered as the hidden geometry of a 3D model identifiable in the configuration of the mesh. Retopologizing thus means replacing the mesh conformation without modifying the apparent geometry of the 3D model.

4 The compositor software are applications that have multiple functions of both modelling and rendering, are used to design a complete 3D scenario or a playable level of video games.

5 We use the normal outgoing to identify the triangular polygon in its visible part, in basic computer graphics the visible part of the mesh is the one with the normal outgoing, the opposite part is transparent instead.
Moving into the Cinema 4D environment, we proceeded to the phases of composition and rendering of 360 panoramic HDRI images in equirectangular format\(^6\) (fig. 5) 12000x6000 px, relying on the OTOY OctaneRender engine to be used in the interactive tour created with Pano2VR. Furthermore, to achieve realistic lighting, the path tracking algorithms used by the rendering engine in question offered excellent performance thanks to HDRI support, mesh emitters and a planetary sun/sky system that simulates the correct inclination of sunlight.

![Fig. 5. Equirectangular image rendered in Cinema 4D with the OctaneRender engine, ready for implementation in Pano2VR.](image)

OctaneRender allows real-time manipulation of lighting in a scene, producing very photorealistic lighting and post-imaging effects.

The Pano2VR Pro software was used for the construction of the virtual tour, in which we connected the various 360° scenarios, defining the exploration routes by means of automatic links, filters, the modification of the main node and the use of integrated maps. The addition of interactive elements such as hotspots, directional audio and video also made the gaming experience more immersive and engaging.

The implementation of Oculus Rift, thanks to gyroscopes and motion sensors, allowed good synchronization with the head movements and a correct positioning of the optical cone, which contributed to another fundamental objective we had set for ourselves, namely the reduction of cybersickness phenomena.

The player chosen to participate in the serious game is VRtourviewer by 3DV.

\(^6\) HDRI (High Dynamic Range Rendering) uses a much higher dynamic range of colours and values than traditional bitmap formats. Instead of encoding colours like a computer monitor (using 24 bits per pixel), the HDRI works on a trichromatic basis like the human eye and has the spectral sensitivity of the human eye, storing the actual luminance values of each pixel. HDR images therefore not only store the colour, but also the intensity and brightness of the light at each point. The range of colours and light that this format can contain is much greater than the RGB scale used in traditional computer graphics.
3.2.3 Questionnaires

Three questionnaires were administered to the players. The first was based on the standardized test sITQ_PQ for measuring "presence" in virtual environments, proposed by Witmer and Singer [45]. It included five questions, evaluated on the basis of the 5-
level Likert scale. It was administered upon completion of the operator phase, before entering the observer phase. The second tested the approval of the experimentation, and included five more rating questions on the 5-level Likert scale; it was administered at the end of the game to all members of the group. The third was aimed at assessing the persistence of knowledge and skills acquired during the experiment, and was administered at school about a week after the VR experience. It included a mixture of Likert-scale, closed and open questions, to investigate a general evaluation of the VR experience and of what was learnt thanks to the VR experience.

3.3. Procedure

During the first year, the development of the project focused on the design of the educational paths. The didactic path was completed from January to April 2019, while the assessment was conducted from May 28th to 30th of the same year. The didactic paths were developed in cooperation with the teachers who joined the experimentation. To maximize the pedagogical value of the didactic path, it was decided to build the activities starting from the didactic plan already scheduled for the 2018-19 school year. The choice, made in agreement with the teachers and school managers, fell on a didactic path centred on the traditional bread of the Puster Valley and the farming tradition, which required a complementary design to allow integration with a process of heritage education.

During the school year, the design of the educational path was divided into 3 phases:
1. introduction of the topic in class (the dolomitic Farmhouse - maso)
2. visit to the site
3. virtual feedback activity

The first phase took place two weeks before the visit to the site, which occurred 4 weeks before the third, experimental phase.

The second phase began with a visit to the Südtiroler Landesmuseum für Volkskunde in Teodone (BZ), where the children followed a path organized by the museum staff, on the rural tradition and the preparation of rye bread in the Puster Valley. The visits of the two classes participating in the experimentation, occurred around 2 months before the third phase in VR.

For the third phase, the children were divided into three groups of seven children (approximately) each, randomly assigned to each group. While one group was operating with VR, the others were involved in activities related to this subject in other rooms, and were not in contact with each other (so that they could not communicate). The VR activity involved each group for one hour. While a group was taking part in the VR activity, each of the seven children was given the chance to take the active role (called "operator", who wore the VR viewer: fig. 8a) and the other six (the "observers") participated by watching the projection and movements of the operator in the virtual environment on a large screen. Each scenario was explored and activated by the operator, while the observers could interact at certain times and could help the operator in case of uncertainty or difficulties in the quizzes.
Fig. 8. The upper image shows a child in the 4th grade of primary school in the role of operator, exploring one of the environments in the serious game. The lower picture shows one moment of collaboration between the group of observers and the operator.
At the end of the experience in one scenario, the operator was replaced by another student in the group, was assessed through the first questionnaire, and then joined the observers. This structure allowed us to consider the game activity as a one-time situated learning activity lasting about 50-60 minutes, where the exposure to immersive visualization lasted 7-8 minutes for each child. In the remaining time, the observers participated in the game through the projection of the operator's action on the screen (see figure 8b).

The observations of the players' behaviour were structured in different forms: the first, for the duration of the experimentation, was a series of video recordings with four cameras, two of which were arranged frontally and laterally with respect to the operator's position, one which was arranged to frame the observer-players and one at 360° to allow a synchronous relationship between the operator, the observers and what was happening during the game in the virtual space. At the same time, a member of the research team recorded each event that was considered significant by the coordinators, so that the audio-video data could be analysed more accurately. After the last child of the group has finished his task as operator, one experimenter debriefed the children by providing them with answers to questions or information about the relevance of the VR device.

3.4. Data Analysis

For this article, the data analysis includes the answers given to the Likert scale questions in the first and second survey as they were (ranging from 1 to 5, 1 being the positive and 5 the negative value). The two Likert scale questions in the third survey were also analysed. Moreover, a total of 13 points could be given for the questions that tested knowledge (see the annex 1), thus creating the variable “Knowledge”. The Class factor represented whether students belonged to the second or fourth grade of the primary school.

The data was analysed using the free statistical software R [46]. First, the 12 variables obtained from those resulting from the first and the second survey, and from the quantitative items in the third survey, were described indicating their frequency of responses. These 12 variables were then correlated using the polychoric correlation, which measures the association between two ordinal variables, under the assumption that each ordinal variable stands for a discrete representation of an underlying continuous variable. The data was then screened through Bartlett’s test to identify the presence of multicollinearity in the data matrix and consequently apply a reduction of the variables. If the Bartlett’s test was significant, a Principal Component Analysis (PCA) would be run to reduce a larger set of variables (which could have a certain amount of correlation) to a smaller set of uncorrelated variables. The Class factor was not included in the first analyses because the number of participants in the two levels was not sufficient to produce reliable evaluations.

Finally, the influence of these 12 variables on learning was evaluated by means of a series of regression models. Each model included Knowledge as a dependent variable, while a set of predictors was obtained by rotating all the combinations of the 12 variables to have a minimum of one to a maximum of three predictors, controlled for the Class factor. The lme4 package for R was used to calculate the regression models
For each predictor, the following parameters are provided in the results: Beta, being the estimated value of the coefficient for the regression; SE, being the standard error (SE) associated with the predictor; the t value, indicating the statistical evaluation; and the probability value (p) associated with the t statistics. The contribution of each effect is computed by excluding the influence of the other predictors. All statistics are considered significant when their associated p-value is below the alpha threshold, set at 0.05.

4. Results

All the Likert-scale items in the surveys showed satisfactory results (they are depicted in Fig. 9). Children expressed an appreciation throughout the various items: negative answers (4 or 5) were less than 3% across the 10 items expressing a judgment on a positive/negative scale.

The series of bivariate polychoric correlations is reported in Table 1. The Bartlett’s test turned out to be significant ($\chi^2(66) = 119.06$, $p < 0.001$); therefore, the reduction of the variables through PCA was appropriate. Using the Kaiser criterion applied to the results of the PCA (“principal” function in R, coupled with an oblimin rotation), five components were extracted with eigenvalues (see “SS loadings” in Table 2) higher than 1 (fit of the model = 0.89). The total variance explained by the five components was 0.71.

Fig 9. The figure describes the frequency of responses to the Likert scale levels, separated for the 12 items. All items ranged from 1 (most positive value, green colour) to 5 or 6 (most negative value, red colour), except the “Useful vs. Learning” item, which ranged from 1 (Useful) to 6 (Enjoying).
Table 1. Polychoric correlations. The Rho values are presented in each cell, representing the relationship between the variable indicated in the row and the column. Negative values indicate an inverse relationship. Significant values are represented in bold, indicating a medium to strong relationship.

<table>
<thead>
<tr>
<th></th>
<th>Involvement VE</th>
<th>Navigation VE</th>
<th>Similarity</th>
<th>Quality VE</th>
<th>Like</th>
<th>Feel exploring</th>
<th>Feel answering</th>
<th>Help learning</th>
<th>Learning through game</th>
<th>Useful vs. Enjoying</th>
<th>Estim. learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed adaptation</td>
<td>0.04</td>
<td>0.67</td>
<td>0.17</td>
<td>0.22</td>
<td>0.56</td>
<td>0.23</td>
<td>0.43</td>
<td>0.05</td>
<td>0.47</td>
<td>-0.18</td>
<td>0.12</td>
</tr>
<tr>
<td>Involvement VE</td>
<td>-0.09</td>
<td>0.19</td>
<td>0.43</td>
<td>0.33</td>
<td>0.53</td>
<td>0.04</td>
<td>0.30</td>
<td>0.24</td>
<td>-0.26</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Navigation VE</td>
<td>-0.04</td>
<td>0.35</td>
<td>0.38</td>
<td>0.04</td>
<td>0.59</td>
<td>-0.17</td>
<td>0.58</td>
<td>-0.26</td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Similarity</td>
<td>0.38</td>
<td>-0.26</td>
<td>0.31</td>
<td>0.18</td>
<td>0.23</td>
<td>0.33</td>
<td>0.33</td>
<td>-0.02</td>
<td>-0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality VE</td>
<td>-0.14</td>
<td>0.50</td>
<td>0.19</td>
<td>0.25</td>
<td>0.49</td>
<td>-0.10</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Like</td>
<td>0.07</td>
<td>0.61</td>
<td>0.36</td>
<td>0.53</td>
<td>-0.59</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feel exploring</td>
<td>0.39</td>
<td></td>
<td>0.41</td>
<td>0.08</td>
<td>-0.51</td>
<td>0.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feel answering</td>
<td>0.30</td>
<td>0.24</td>
<td>0.26</td>
<td>0.53</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Help learning</td>
<td>0.05</td>
<td>0.03</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning through game</td>
<td>-0.49</td>
<td>0.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Useful vs. Enjoying</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.26</td>
<td></td>
</tr>
</tbody>
</table>

The first component includes Like, Feel answering, Speed adaptation and Navigation VE: this means that these four variables are strongly associated with each other. A possible interpretation could consider that a positive feeling when answering the questions could depend on rapid adaptation to the VE and ease of navigating in the VE. The second component again includes Navigation VE, as well as Quality VE and Learning through game: navigation may positively influence the perceived quality of the vision (that did not interfere with the task) and, in turn, the evaluation that learning is possible through a game. The third component shows that the more positive the evaluation of the exploration is, and the (higher) the degree of involvement in the VE, then the closer this experience is to enjoyment rather than usefulness. The fourth component indicates that the more the students evaluated the VE to be similar to the real environment, the more they thought they had learned from it. The fifth component is defined by one single variable: the estimation of how significantly this experience can help learning.
Table 2. Results of the PCA. The rows in italics present the initial variables; columns PC1 to PC5 show the variables’ loads on the five PCs; h2=communalities, com=complexity. A negative value indicates an inverse relationship with respect to the component and the other(s) variable(s).

<table>
<thead>
<tr>
<th>Variable</th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
<th>PC5</th>
<th>h2</th>
<th>com</th>
</tr>
</thead>
<tbody>
<tr>
<td>Like</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feel answering</td>
<td>0.64</td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>Speed adaptation</td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>Navigation VE</td>
<td>0.56</td>
<td>0.56</td>
<td></td>
<td></td>
<td>0.80</td>
<td></td>
<td>2.3</td>
</tr>
<tr>
<td>Learning through game</td>
<td></td>
<td></td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
<td>1.1</td>
</tr>
<tr>
<td>Quality VE</td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.5</td>
</tr>
<tr>
<td>Feel exploring</td>
<td></td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td>0.75</td>
<td>1.0</td>
</tr>
<tr>
<td>Involvement VE</td>
<td></td>
<td></td>
<td>0.69</td>
<td></td>
<td></td>
<td>0.52</td>
<td>1.2</td>
</tr>
<tr>
<td>Useful vs. Enjoying</td>
<td></td>
<td></td>
<td></td>
<td>-0.66</td>
<td></td>
<td>0.68</td>
<td>2.1</td>
</tr>
<tr>
<td>Estim. learning</td>
<td></td>
<td></td>
<td></td>
<td>-0.81</td>
<td></td>
<td>0.85</td>
<td>1.5</td>
</tr>
<tr>
<td>Similarity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.72</td>
<td></td>
<td>1.7</td>
</tr>
<tr>
<td>Help learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.94</td>
</tr>
</tbody>
</table>

The regression models regarding the children’s Knowledge (the amount of correct answers in the verification one week after the experience) converged for all combinations of predictors, but only a small number of them reached the significance level. Table 3 reports four models, the only ones in which one or more predictors were significant. The first model showed the key effect of Quality VE, indicating a direct relationship between the quality of the VE, which did not interfere with the task, and the knowledge displayed by the children. The predictor Like approached the significant threshold but a concrete relationship cannot be considered reliable, given that 34 children out of 36 indicated the most positive value in the scale (and a real trend cannot be estimated). The second model (Table 3, section a) confirmed the key effect of Quality VE and, additionally, revealed an interaction between Class and Similarity (Figure 10a): as the perceived similarity of the VE with reality increased, the knowledge performance either decreased (for children in the second grade) or increased (for children in the fourth grade). The third model showed the interaction between Involvement VE and Class being marginally significant. Figure 10b shows that knowledge performance decreased as the perceived involvement in the second grade
increased, while the opposite relationship approached the significant level for the fourth grade. In the last model, a difference caused by Class was noticed in the relationship between Useful vs. Enjoying and Knowledge. Figure 10c shows that the best performance was obtained in the second grade by the children who considered the VE experience more enjoyable, while in the fourth grade by those who considered the VE experience being more useful.

Table 3: Regression models: the results obtained from the regressive models. Beta (SE) are the estimated values for the relationship and, within brackets, the associated Standard Errors; t value refers to the statistics computed for the predictor; p value is the probability associated with the statistics; dof = degrees of freedom.

| Formula: \( \text{lm(Knowledge ~ QualityVE + Like + Estim. Learning, data)} \) |
|---------------------------------|--------|--------|--------|
|                                | Beta (SE) | t value | p value |
| (Intercept)                     | 12.633 (1.721) | 7.340 | 0.001 |
| Quality VE                     | -1.074 (0.500) | -2.148 | 0.041 |
| Like                             | -1.573 (0.856) | -1.838 | 0.078 |
| Estim. learning                | -0.486 (0.495) | -0.981 | 0.336 |
| Residual SE: 2.615 (26 dof)    |          |        |        |
| Multiple R²: 0.245             |          |        |        |
| Adjusted R²: 0.158             |          |        |        |

a) Formula: \( \text{lm(Knowledge ~ Similarity*Class + QualityVE*Class, data)} \)

<table>
<thead>
<tr>
<th></th>
<th>Beta (SE)</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>8.215 (1.749)</td>
<td>4.698</td>
<td>0.001</td>
</tr>
<tr>
<td>Similarity</td>
<td>1.629 (0.900)</td>
<td>1.810</td>
<td>0.081</td>
</tr>
<tr>
<td>Class4</td>
<td>3.503 (2.664)</td>
<td>1.315</td>
<td>0.199</td>
</tr>
<tr>
<td>QualityVE</td>
<td>-1.629 (0.697)</td>
<td>-2.335</td>
<td>0.027</td>
</tr>
<tr>
<td>Similarity:Class4</td>
<td>-3.717 (1.309)</td>
<td>-2.839</td>
<td>0.008</td>
</tr>
<tr>
<td>QualityVE:Class4</td>
<td>0.975 (1.012)</td>
<td>0.963</td>
<td>0.344</td>
</tr>
<tr>
<td>Residual SE: 2.694 (29 dof)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple R²: 0.343</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R²: 0.229</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b) Formula: \( \text{lm(Knowledge ~ InvolvementVE*Class, data)} \)

<table>
<thead>
<tr>
<th></th>
<th>Beta (SE)</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>5.861 (1.852)</td>
<td>3.165</td>
<td>0.003</td>
</tr>
<tr>
<td>InvolvementVE</td>
<td>2.030 (1.332)</td>
<td>1.525</td>
<td>0.138</td>
</tr>
</tbody>
</table>
Figure 10: The relationship between the dependent variable (Knowledge, on the y-axis) and some predictors (on the x-axes), separated for the two-level Class factor (red line: second grade; cyan line: fourth grade) is depicted in the three graphs: a) Similarity; b) Involvement in the VE; c) Usefulness (values toward the lowest end) versus Enjoying (values toward the highest end). The lines represent the estimated relationship, while the grey ribbons represent the estimated variability of the relationship.

5. Discussion

The aim of the research was to verify the effectiveness for learning of a didactic path, organized as a serious game, to be carried out in an immersive digital environment, the theme of which was the cultural heritage and which followed a project methodology
that kept the welfare of the child at the centre of its concerns (User Centered Design: UCD).

The main results could be summarized as follows: a) all children expressed their appreciation with respect to the activity, b) their evaluations revolve around four main parameters, c) the relationship between VE and students’ learning is mediated by few features and age and therefore d) the didactic procedure seems to be effective in developing a successful knowledge experience.

The majority of children expressed their appreciation for the activity in all the collected items. This result is consistent with those reported in the literature [48], in which the VR methodology was reported to attract the interest and attention of the students.

To some extent, it is quite normal that an activity engaged in outside the School, which also includes digital technology, would capture the students’ appreciation and interest [49]. However, it must be considered that, in our design, students were exposed to the VE for only a few minutes and, therefore, had it not been satisfying, they could have expressed their disappointment for a broken promise. In the experience however, each student was personally involved with the game about for 50-60 minutes (for less than 10 minutes with the device as the operator, and for about 40-50 minutes in the group as an observer), and in the remaining three hours they were involved either in a debriefing (right after the group completed the tasks) or in other activities (while other groups were using the device). Given that each child in the group experienced the role of both “operator” and “observer”, the experience was shared among all the children. Therefore, no child felt uneasy in being alone in a virtual space for a long time7, and in fact - considering the results of the questionnaires following the experimentation - they were likely to have experienced the game in a more continuous manner.

The principal component analysis showed that only four parameters could be considered as relevant to the investigation of the students’ evaluation of the VE. The first component is related to rapid adaptation and navigation in the VE, and to the feeling while executing the task in the VE. This component is expected to merge technical issues and developments with the human performance. The second component also included navigation and, additionally, the quality of the VE and the gamification: it may represent the need for better scenarios, to improve the relationship between gamification and learning. The third component merged exploration, involvement and enjoyment: it could represent how playful the environment can be. The fourth relevant aspect to be investigated connects similarity and the student’s estimation of her learning through the activity. That is, the less difference a student finds between VE and RE, the more she predicts that she will learn through the activity. Clearly, the quality of the representation in the VE contributes decisively to this aspect. In the literature, ambiguous interpretations are reported. On the one hand, some scholars support an orientation towards high fidelity scenarios [50][51] considering sound as well [52][53][54]. Others state that sophisticated scenarios are not necessary for the point of view relevant to learning [55][56][57]. Based on results obtained by the serious game proposed in our study, the realism of the scenes could contribute both to

7 The exploration and solution of each level of the game required an average time ranging between 8 and 10 minutes which, even in the absence of experimental data, can be considered an effective timeframe to eliminate the risks related to cyber sickness due to overexposure of the children to VR. The entire game takes between 50 and 60 minutes to complete.
the similarity between the experience in VR and that in the Museum of Teodone, and to the contextualization of the activities carried out in VR with real life. This statement is supported by the high percentage of serious games designed for children of younger ages that use non-realistic 2D graphics [9].

A fifth independent component was represented by a single variable, gauging how much this activity is expected to help learning, which is not associated with any of the other variables. It may be the case that students responded randomly to this question, which is more related to the evaluation of teaching methods than of the activity itself. Therefore, this component would deserve further investigation, to determine whether the question or its content were creating a bias or, alternatively, if this was really an independent concept.

In general, the sample in the present study is not very large and, given that the stability of results from the PCA would suffer from the size of the dataset, the results need to be supported by further studies, with larger datasets and different environments. An extension of the sample is planned through the experimentation of our immersive didactic paths in other schools of the same territory.

The regressive models indicated that, regardless of the age group, the perceived quality of the VE is an important factor to foster learning: the greater the quality seen in the VE, the higher the students’ knowledge in correctly answering the questions. Nevertheless, there were differences in the evaluation provided by the two age groups in other items. The younger students, aged 7-8, were more attracted by the playful quality of the VE, rather than its utility, while the opposite was true for the older students, aged 9-10. This result impacted the students’ performance on the questions related to their learning: the more correct responses they provided, the more the second grade and fourth grade students evaluated their experience as being enjoyable or useful, respectively. Other differences in the knowledge between the age groups were explained by the perceived similarity of VE with respect to a real environment, and by the level of involvement produced by the VE. While these two evaluations increased, the knowledge of the fourth-grade students increased and for those in second grade decreased. Presumably, at age 7-8, the children who are skilled enough to process the VE situation are much more attracted than those who are not, and, in turn, the former are more distracted from the real goal of the experience (that is, learning) than the latter. At ages 9-10, children are more likely to discern what is relevant and what is secondary for the given goal. They could have appreciated the high level of reality in the VE, in order to embed themselves in the environment and think about it, to focus on what could be relevant for the farmers in the Farmhouse. Other items did not show any relationship with Knowledge and could be considered not relevant.

Both the four main components and the regression models are opening an important methodological discussion on the associations that exist between technical and user aspects. Developers of scenarios need to know what technical improvement is desirable to work on, in order to influence user experience. According to the present investigation, students in primary school did not consider human performance and technical aspects as separate categories. The implicit evaluation (namely, the knowledge associated with students’ learning) was predicted by their explicit evaluation of interactive issues (such as similarity to RE and perceived involvement) and of personal assessment (whether relevant for enjoyment or utility). Rather, they intertwined their performance and learning with many parameters perceived from the
VE. This could be considered an unexpected outcome, given that such young users were not thought capable of evaluating technical aspects. It is recommended to take into account the effectiveness of the visualization of the VE, especially by "calibrating" the representative choices (model detail, texturization, lighting, locomotion and binocular/monocular projection) for the specific users of the experimentation or implementation.

The findings from this study indicate that learning while using VE may be enhanced when the age of the students is higher than nine, while younger students are more attracted by the “game” and are not exploiting this opportunity for learning. The approach is not new [58], but the added-value of the present work lies in the continuity of the research path. It started with the pedagogical project, continued with the design of the game and the architectural environments, then with the engineering and the constitution of the representation device, and ended with the experimentation. The developmental reason for this difference needs to be elucidated and deserves further investigation, as it lies beyond the purposes of the present study. However, cognitive development could help to understand why some technical aspects are linked to certain personal evaluations. Future studies are expected to contribute to untangling these relevant aspects.

In conclusion, the present study showed that VE could be a proficient tool for learning from a certain age on. Moreover, it highlighted the relevance of some key features of VE, while creating scenarios to be used for learning by children in primary school. Although there are some limitations⁸, this study has explored VE in an ecological context and is opening a research field combining education and documentation of heritage through digital surveying and three-dimensional digitization.

Acknowledgments. The VAR.HEE. project – Virtual and Augmented Reality for Heritage and art Education in school and museum Experiences – was funded by the Free University of Bozen/Bolzano with a call for proposals from the Central Research Commission in 2017. The project lasts three years, started in January 2018 and will end in December 2020 (June 2021, after Covid-19 health emergency).

Despite the work has been developed in team, paragraphs 1, 3.1, 3.2.1, 3.2.3 and 3.3 have been drafted by Alessandro Luigini, paragraph 2 by Monica Parricchi, paragraph 3.2.2 by Alessandro Basso, paragraphs 3.4 and 4 by Demis Basso. Paragraph 5 was drafted jointly by Alessandro Luigini, Monica Parricchi and Demis Basso.

The teachers who participated in the co-design phase are Rita Martelli, Susanna Saporri, and Emanuela Turatto, and in the experimentation phase have collaborated Starlight Vattano, Valentina Dematte and Sara Pellegrini.

⁸ The limitation of the sample is being overcome through the testing of the educational path in collaboration with other schools of the Province of Bozen, expanding the sample many times. The additional limitations introduced by the Covid-19 health emergency have limited the possibility of implementing the work plan planned for 2020. The development of the project is now focused on the transposition for remote experience of the immersive pathways. In any case, a new experiment with over 200 primary school students is planned.
References