

Passive haptics and conversational avatars for interacting with ancient Egypt remains in high-fidelity virtual reality experiences

*Original*

Passive haptics and conversational avatars for interacting with ancient Egypt remains in high-fidelity virtual reality experiences / Cannavo', A., Pacchiotti, S., Retta, N., Terzoli, M., Spallone, R., Lamberti, F.. - In: ACM JOURNAL ON COMPUTING AND CULTURAL HERITAGE. - ISSN 1556-4673. - ELETTRONICO. - 17:2(2024), pp. 1-28. [10.1145/3648003]

*Availability:*

This version is available at: 11583/2985063 since: 2024-05-08T11:21:46Z

*Publisher:*

ACM

*Published*

DOI:10.1145/3648003

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

ACM postprint/Author's Accepted Manuscript

© ACM 2024. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in ACM JOURNAL ON COMPUTING AND CULTURAL HERITAGE, <http://dx.doi.org/10.1145/3648003>.

(Article begins on next page)



# Passive Haptics and Conversational Avatars for Interacting with Ancient Egypt Remains in High-Fidelity Virtual Reality Experiences

ALBERTO CANNAVÒ, SIMONA PACCHIOTTI, and NICOLA RETTA,

Politecnico di Torino, Dip. di Automatica e Informatica, Turin, Italy

MARTINA TERZOLI, Fondazione Museo delle Antichità Egizie di Torino, Dip. Collezione e Ricerca, Turin, Italy

ROBERTA SPALLONE, Politecnico di Torino, Dip. di Architettura e Design, Turin, Italy

FABRIZIO LAMBERTI, Politecnico di Torino, Dip. di Automatica e Informatica, Turin, Italy

As Extended Reality (XR) continues to grow, new possibilities arise to provide users with novel ways to experience cultural heritage (CH). In particular, applications based on Virtual Reality (VR) like, e.g., virtual museums, have gained increasing popularity, since they can offer new ways for preserving and presenting CH content that are not feasible in physical museums. Despite the numerous benefits, the level of immersion and presence provided by VR experiences still present challenges that could hinder the effectiveness of this technology in the CH context. In this perspective, it is crucial to provide the users with high-fidelity experiences, in which also the interaction with the objects and the characters populating virtual environments are realistic and natural. This paper focuses on this challenge and specifically investigates how the combined use of tangible and speech interfaces can help to improve the overall experience. To this aim, a immersive VR experience is proposed, which allows the users to manipulate virtual objects belonging to a museum collection (in the specific case, Ancient Egypt remains) by physically operating on 3D printed replicas and to talk with a curator's avatar to get explanations by using their voice. A user study was conducted to evaluate the impact of the considered interfaces on immersion, presence, user experience, usability, and intention to visit, comparing the richest configuration against simpler setups obtained by either removing the tangible interface, the speech interface or both (and using only handheld controllers). The results showed that the combined use of the two interfaces can effectively contribute at making the CH experience in VR more engaging.

CCS Concepts: • **Human-centered computing** → **Virtual reality; Haptic devices; Natural language interfaces; Applied computing** → **Arts and humanities**.

Additional Key Words and Phrases: Virtual Museums, Virtual Humans, Tangible User Interfaces, Speech User Interfaces

## 1 INTRODUCTION

Over the last years, important advancements in the field of eXtended Reality (XR) and the consequent possibility to devise new interaction paradigms by leveraging this technology have attracted more and more the attention of researchers from different application domains. In particular, the adoption of XR has proven capable of bringing many benefits in the cultural heritage (CH) domain, e.g., by improving accessibility of cultural heritage sites or profiling visitors to offer them tailored content [23, 34]. Among the numberless application scenarios of this

---

Authors' addresses: Alberto Cannavò, [alberto.cannavo@polito.it](mailto:alberto.cannavo@polito.it); Simona Pacchiotti, [simona.pacchiotti@studenti.polito.it](mailto:simona.pacchiotti@studenti.polito.it); Nicola Retta, [simona.pacchiotti@studenti.polito.it](mailto:simona.pacchiotti@studenti.polito.it),

Politecnico di Torino, Dip. di Automatica e Informatica, C.so Castelfidardo, 34/d, Turin, Italy, 10138; Martina Terzoli, Fondazione Museo delle Antichità Egizie di Torino, Dip. Collezione e Ricerca, Turin, Italy; Roberta Spallone, Politecnico di Torino, Dip. di Architettura e Design, , Turin, Italy, [roberta.spallone@polito.it](mailto:roberta.spallone@polito.it); Fabrizio Lamberti, [fabrizio.lamberti@polito.it](mailto:fabrizio.lamberti@polito.it), Politecnico di Torino, Dip. di Automatica e Informatica, C.so Duca Degli Abruzzi 24, Turin, Italy, 10129.

---

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2024 Copyright held by the owner/author(s).

ACM 1556-4673/2024/2-ART

<https://doi.org/10.1145/3648003>

technology, this work mainly focuses on preserving and presenting CH content [32]. As a matter of example, Augmented Reality (AR) technology and holographic displays have been successfully applied to digital culture heritage in order to aesthetically enhance exhibits and heritage sites [47], provide virtual guided tours [28], and support the learning of ancient languages and culture through novel gamified approaches [50]. Even looking at Virtual Reality (VR) technology, it is possible to notice that the number of works that exploit immersive technologies and leverage novel human-computer interaction (HCI) techniques for exploration and education purposes is progressively growing in the CH literature [4, 58, 75]. Specifically, VR experiences such as virtual museums, have gained increasing popularity.

In fact, virtual museums offer the opportunity to complement traditional, physical visits of the collections, letting users enjoy them from anywhere and at anytime [32]. They also enable the implementation of alternative visiting modalities that are not feasible in physical museums: for instance, users can be allowed to examine and interact with 3D replicas which could represent not only collection objects that are fully preserved, but also objects that are partially preserved, damaged, or currently not available for display (e.g., because loaned-out) [32]. Given the above opportunities, virtual museums are recognized as one of the most cost-effective and dynamic means of connecting visitors to a CH environment, its artifacts, and the associated knowledge [42].

Despite the numerous benefits, open challenges still remain regarding the level of immersion and presence that can be delivered through these VR applications [55]. The final goal of VR should be to fully immerse the users in the virtual environment by making them experience the same physical and psychological reactions to the provided stimuli that they would feel in the real world [2] (i.e. sense of immersion); in this way, the illusion of being part of that experience (i.e., sense of presence) [62] would be fostered too. Therefore, to maximize the effectiveness of VR use – and, in particular, enhance the perceptual, cognitive, and communicative potential of CH content [6] – it is crucial to provide high-fidelity, interactive experiences that extensively stimulate the human sensory system [49]. In this respect, it is worth observing that VR experiences that envisage limited interactions with the surroundings or leverage only inanimate virtual environments can be expected to negatively impact users' engagement [74]. For this reason, significant research activities are being devoted to improve the way in which users can interact in VR [46, 64], as well as to endow these experiences with virtual humans (VHs) designed to closely resemble the appearance and behavior of real individuals [42]. Studies in the literature already reported psychological effects due to the design of VHs and available interactions [16].

For what it concerns the interaction with the VR experience and its elements, researchers have increasingly focused their attention on embodied and tangible HCI [30]. In particular, the literature shows that Tangible User Interfaces (TUIs) can be used to provide accurate feedback for different physical shapes and materials [9, 61], and can also enable new paradigms for I/O and interaction with digital information [35]; this is possible since they allow users to physically engage with this information, by literally letting them grasp and manipulate it with their hands [60]. Incorporating haptic feedback can also improve the user experience in VR, enhancing the sense of immersion [54]. Focusing on CH, TUIs can play a fundamental role in the current museology trend which aims to let the users physically interact with museum collections using touch [31]. The idea is that, in this way, the users' senses can be stimulated more, thus improving the learning experience and engagement [63]. Although touching the objects can be beneficial from the mentioned perspectives, most of the times is actually impractical. It can expose the objects to contaminants from the users' skin, leading to potential harm to the collection; moreover, the overuse, accidental dropping, or scratching of the objects can result in their structural damage [52]. In addition to the risks for the objects, concerns for the users should also be considered, as some objects may pose risks due to their weight, sharp edges, or the presence of hazardous substances [53]. The use of TUIs could allow to overcome these issues. There are basically two approaches to implement TUIs, i.e. active or passive, though hybrid approaches are also possible [56]. Active approaches can be characterized by different levels of complexity. The most common example is represented by handheld controllers: although these devices support vibrotactile feedback and currently represent the standard interface to interact with objects

in VR, they are not suitable for fully stimulating the sense of touch and to provide, e.g., information about the objects' weight, shape, and texture [21]. Other active approaches leverage motors, electromagnets and further mechanisms to exert forces on the user's body: however, the higher realism of these stimuli is usually paid with complex and sophisticated setups, which may affect usability [45]. Passive approaches, in turn, typically referred to as passive haptic interfaces, or simply passive haptics (PHs), rely on physical props not endowed with any sensor or actuator, which are connected to virtual objects and replace them for VR interaction. PHs are similar to their virtual counterparts in terms of relevant haptic properties, and can offer a deeply realistic perception of physical characteristics, since touching real surfaces eliminates the need to simulate properties such as texture, hardness, weight, shape, and size, thus increasing the sense of immersion [45].

For what it concerns the second major aspect being investigated to enhance the user experience in VR, i.e. the use of VHs, by considering relevant literature it can be observed that they typically range from pure decorative elements to intelligent agents supporting the users in different ways [12]. Focusing on the CH domain, over the last years VHs in the form of virtual guides started to be embedded in many VR experiences, as they proved to be an effective way to convey information by engaging users in an interactive exchange of knowledge that can promote participation and attention, leading to a deeper understanding [19, 24]. In fact, VHs can make the stories told in virtual environments more believable, can influence the users in a positive and constructive way [38] by motivating them to enjoy the content longer, and can enable unstructured narrative experiences without losing critical information [70]. However, the introduction of VHs in VR does not come without challenges [38], since the lack of consistency in the replica's realism can lead the users to experience unintended cold or eerie feelings [41]. This effect is known as the "uncanny valley" [44] and it is the result of a perceptual mismatch due to the observation of conflicting cues in the avatar's appearance, e.g., unnaturally large eyes in a realistic and well-proportionated face [57], etc. To limit this effect, research proposes mechanisms such as increasing the character's physical attractiveness, avoiding altering the natural body structures and proportions, considering not only the gender but also the details of body characteristics (e.g., skin color and unique features) when designing VHs that have to resemble real people, and providing time for the users to get accustomed to the VHs [57]. Realism includes the possibility for the users to interact with VHs naturally, to avoid affecting the sense of presence and introducing other possible mismatches and inconsistencies [17].

To this aim, several works have proposed to use Speech User Interfaces (SUIs) [22], based on speech recognition and synthesis [22, 59]. The use of voice in HCI has already been proven to be capable of enabling plausible 3D experiences [13] and stimulating the communication of CH content [8].

Moving from the above considerations, this paper presents the design and development of a high-fidelity, curated experience in VR that supports interaction with CH content based on both TUIs and SUIs. The paper builds upon a previous work [55] that proposed an experience in which a virtual curator guides the users in the exploration of Ancient Egypt remains. The work involved experts from the Museo Egizio in Turin, Italy, who provided historical background and participated in the generation of the VR assets. With respect to the previous work, in which interaction with the VH and the objects was mediated by the handheld controllers, the experience designed in the present paper moves some further steps towards the adoption of more natural interaction modalities. In particular, the remains can be physically manipulated as they are managed in the VR experience as passive haptics; furthermore, the users can ask for information about the remains through their voice by interacting with a conversational avatar representing the curator.

A user study has been carried out to assess the effectiveness of integrating TUIs and SUIs in a CH experience from different perspectives. More specifically, the proposed experience has been compared with that in [55], with the aim of studying the impact that the considered interfaces can have on users' engagement. Engagement has been investigated by collecting data regarding the perceived sense of immersion, presence, user experience, usability, and intention to visit. In order to isolate the contribution brought by each interface, the study was carried out in the form of a breakdown analysis.

## 2 RELATED WORK

As said, the goal of the proposed experience is to combine TUIs and SUIs for letting the users physically interact with remains and with a VH representing a museum's curator into a VR environment. To the best of the authors' knowledge, there are no examples of VR applications that integrate in the same experiences these two interfaces. Advantages could be expected in terms of realism and naturalness of the interaction, with positive effects on the overall engagement. In the remaining of this section, the above fields will be reviewed, with the goal of extracting helpful hints for the design of the experience.

### 2.1 Tangible Interaction

The use of tangible interfaces to improve the realism of virtual CH experiences is not new. As a matter of example, the work in [51] proposed a system named VIRTual EXhibit (VIRTEX), which supports interactive experiences with physical replicas of real artifacts by means of TUIs created using 3D printed, touch-sensitive props. During the experience, the users can touch and manipulate the replicas to retrieve information about the artifacts. Touch sensors have been added to the surface of the replicas in specific areas of interest. When the users touch one of these areas, the playback of a short video or animation dealing with the history of the artifact or the characteristics of the specific area is triggered. The content is presented on a detached visualization area, i.e. a screen placed next to the replicas or a wall projection. A sensor is also used to track the replicas in order to correctly visualize the corresponding virtual objects. The system was successfully employed at the Provincial Museum of Archeology in Ename, Belgium to exhibit both artifacts with a real scale and larger objects that could not be picked up, such as monuments or sites. The choice to detach the visualization of the virtual objects from their physical counterparts make it easy to deliver the experience to both single and multiple users. However, this separation could impact usability, negatively affecting the users' engagement. For this reason, the present paper adopts immersive VR technology, and uses PHs to tightly couple the visualization of the virtual objects to the physical props that represent them.

One of the first studies that used PHs in VR is reported in [29]. The work investigated the impact that physically touching objects can have on the way the users perceive the realism of a virtual environment. Although rather dated now, the work laid the foundations for research in this field, since it empirically demonstrated the validity of using TUIs to provide feedback regarding textures and tactile stimuli of the surrounding environment. Considering more recent literature, the work in [46] proposed a VR system that allows the users to interact with PHs of CH artifacts realized through 3D printing. The system featured a hand-tracking module that allowed the users to touch the physical replicas by using their hands. As reported by the authors, one of the limitations of their system regarded the accuracy achieved by the hand-tracking technology. Because of occlusions, the tracking of objects deteriorates rapidly when the users grab the virtual objects with their hands. For this reason, it was not possible to grasp an object directly, and the controller was used as a handle for the object. The controller was used also to activate other system functionalities, e.g., to change the visualization and modify the appearance of the virtual objects (e.g., painting on their surface). The results of the performed experiments demonstrated that the VR experience could benefit from a more direct interaction, which would increase users' involvement. Hence, in the present paper, the use of a direct (i.e., non mediated by the controllers) interaction with the PH is investigated, leveraging the hand-tracking capabilities of the selected VR headset. Another cue coming from this study regards the need to make the virtual and real objects as much similar as possible, reproducing not only the shape but also other physical characteristics like, e.g., the weight and its distribution; in this way, the experience would provide a more realistic tactile stimulation and, expectfully, a higher engagement.

Previous studies have investigated how far the discrepancy between real objects and their digital counterparts can be pushed in a VR experiences before breaking the illusion, with negative effects on the user experience. For instance, results reported in [61] showed that the more the physical object used to control the virtual object is

similar to it, the better is the experience. The work in [64] specifically investigated these aspects in the context of CH, by leveraging a VR experience in which the users had the possibility to interact with museum objects in two modalities: by manipulating 3D printed replicas of the objects, or by manipulating physical acrylic boxes hosting the replicas. Although the system did not support hand tracking (hence, virtual representations of the users' hands were not displayed in the virtual environment), the users could feel the surface and the shape of the objects by touching the printed replicas; this was not possible when using the acrylic box. The box was equipped with an HTC Vive Tracker to track its movements and reconstruct them. The results of a user study showed that the possibility to touch and feel the surface of the 3D printed objects made the participants more engaged than when using the box. Similarly to [46], the authors argued that further investigations were needed to explore the possibility of offering tangible experiences in which other physical characteristics of the objects (weight, materials, etc.) are preserved also in the virtual environment. Hence, in the present paper particular attention was put on making the PH reproduce not only the shape, but also the surface and weight of the real remains.

## 2.2 Virtual Guides and Speech Interaction

The literature includes already a significant number of studies that investigated different aspects related to the use of VHS in CH experiences [1, 42]. For instance, the work in [11] reports on a study aimed to compare three storytelling approaches to support the users in visiting a virtual art gallery. The considered approaches were based on: traditional panels, where descriptions of the artworks were shown as text; a narrating voice, which simulated traditional museum audio guides; a VH guiding the users throughout the virtual environment. The results of the study showed that, compared to the traditional panel- and audio guide-based approaches, the VH was able to increase the users' attention and involvement, thus contributing to better content delivery and learning. The authors of [38] focused on the realism requested by VHS in CH applications. In particular, they proposed a cost-effective methodology for creating high-fidelity virtual guides. The methodology encompassed the use of commercial tools for modeling VHS and leveraged a motion capture suite for recording full-body animations, thus ensuring expressive body moves and gestures. Particular attention was devoted to face animations, since they allow the VH to expressively tell the story presented in the CH experience. In this case, the authors exploited an approach based on recorded audio clips and the use of a software for the automatic generation of lip-sync animations; this approach was preferred to traditional motion capture not to limit the flexibility of the proposed methodology. A user study validated the hypothesis that the devised methodology could generate effective VHS, which were perceived as life-like and realistic by the users when embedded in XR experiences. The study in [67] investigated the impact that three types of VHS, i.e. a museum curator, a security guard, and a museum visitor, may have on the credibility of the narration and the emotions they can arouse in a virtual CH experience. The results indicate that the VH representing a museum visitor was able to elicit stronger empathy and emotional involvement in the users than the other types of VHS, but the virtual curator acting as a specialist conveyed admiration and appreciation. In an extension of this work [68], the authors highlighted the need to align and fine-tune the narrative styles and the presented CH content to the role interpreted by the VHS; moreover, they underlined the importance of the affective component in the storytelling. Based on the findings of the above studies, in the present work a highly realistic avatar of a real curator was created using photogrammetry; the avatar accompanies the users in the experience by providing instructions and explanations. For the sake of flexibility, the avatar was animated using automatic tools, but voice recordings were used to better convey emotions.

Regarding the role that virtual guides should have in a virtual CH experience, the work in [5] studied whether the users prefer a free exploration of an exhibit environment accompanied by a VH who presents them the a given content when they explicitly show their interest in it, or be guided by a VH that leads them on a predefined path. Experimental results yielded no clear preferences for either role. Therefore, the authors suggested to combine

the two conditions in order to merge the benefits and ensure higher user acceptance. In the present work, this approach was followed, having the virtual curator introducing the content of the experience, then letting the users choose the explanations they are interested in, but providing them hints to make sure they do not lose any part of the experience.

Regarding the ways to communicate with VHS, interfaces based on voice commands and speech recognition have long been of great interest, since they are considered the most natural way to support HCI [22]. The possibility to use the voice to support CH experiences is not new, as the first examples leveraging voice commands date back to the early 2000s [65]. The use of natural language, however, is more recent. For instance, this kind of interaction has been used in the work reported in [59], whose goal was to transform a traditional museum into a so-called “vocal museum”. To this aim, the authors proposed a system combining technologies pertaining Internet of Things (for indoor localization) and Artificial Intelligence (for implementing a chatbot) to make the users experience a custom visit of the museum. Once the user is localized in front of certain artifact, he or she can ask for (and receive) information about it by interacting using the voice or text messages via a conversational interface based on natural language. Despite the promising benefits deriving from the use of this interaction method, the work lacks a talking avatar to communicate with, and its effectiveness has not been validated in VR (as content is presented/displayed on the screen of a mobile device).

Nevertheless, VHS capable of a bidirectional speech-based communication are regarded as social entities able to make the users react to them similarly to how they would do with real people [25]. Thus, SUIs are considered as particularly important for the development of virtual CH experiences incorporating VHS. For instance, the work in [22] presented a VR application aimed to let users with motor disabilities visit a museum exhibit. The users could ask a virtual curator to provide information about the artifacts by leveraging voice commands. To improve the flexibility of the system in recognizing different inputs, voice commands issued by the users are automatically augmented through a semantic mechanism capable of inferring linked concepts. In [66], a system was proposed in which two life-sized and photo-realistic characters play the role of virtual guides and interact with the users by means of natural language. The two VHS were embedded with animations to perform gestures and other forms of non-verbal communication. The virtual guides could answer general questions regarding the Museum of Science in Boston, USA and suggest the users to check out specific content based on their expressed interest. The results of a user study confirmed the feasibility of using natural language to interact with the virtual guides and its effectiveness in fostering engagement. However, the authors pointed out the need to integrate facial expressions and eye gaze, as well as including idle animations (when users are not interacting with the VHS) to improve the experience. Based on the above outcomes, in the present work the avatar representing the curator has been endowed with the ability to communicate in natural language on a set of topics related to the experience. Moreover, the avatar has been provided with the ability to gaze the user and with automatically generated face animations.

### 3 MATERIAL AND METHODS

As illustrated in Section 1, the VR experience illustrated in the present paper enables a high-fidelity interaction with a virtual curator and the replicas of handheld objects belonging to a museum collection (in the specific case, the Museo Egizio) by making use of natural HCI modalities. The architecture supporting this experience (illustrated in Figure 1) includes a number of components, which are described in detail in the following sub-sections.

#### 3.1 VR Systems

The core of the architecture is represented by a Unity<sup>1</sup> application, which implements the required logic and integrates all the components that are needed for letting the users participate in the VR experience. Two VR

<sup>1</sup>Unity: <https://unity.com/>

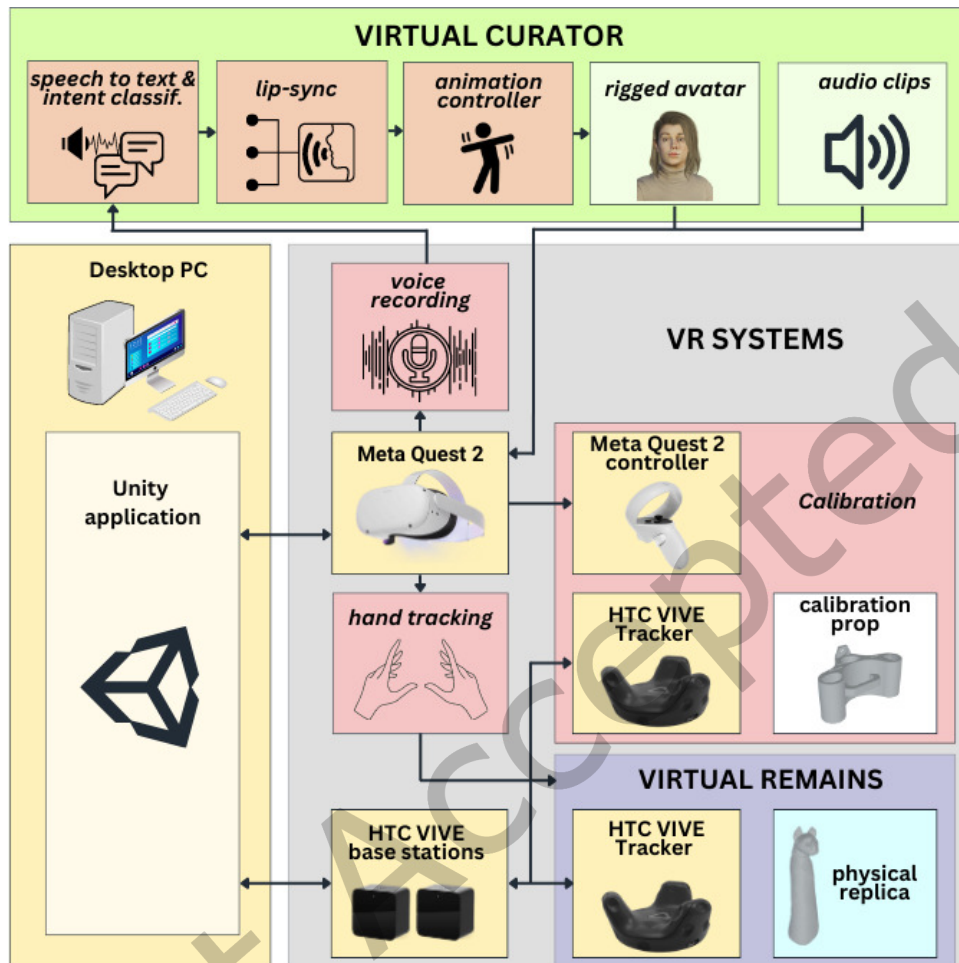


Fig. 1. Architecture supporting the proposed VR experience.

systems are leveraged to support different functionalities of the experience, i.e. the HTC Vive and the Meta Quest 2. The HTC Vive system is used to track the PHs; to this purpose, HTC Vive Trackers attached to the physical replicas are used (more details will be provided in Section 3.3). The Meta Quest 2<sup>2</sup> headset, in turn, is used to make the user visualize the avatar of the curator, the remains, and the virtual environment. The hand tracking capabilities of the headset are used to enable hands-free interaction via the Unity XR Interaction Toolkit<sup>3</sup> (the Meta Quest 2 headset was preferred to the HTC Vive one because of its experimentally observed superior performance, especially in handling occlusions caused by object manipulation); integrated speakers and microphones are leveraged to support speech interaction with the curator. The Meta Quest 2 handheld controllers are used for calibration purposes, i.e. to align the reference system of the headset with that used by the PH. The calibration consists in placing a HTC Vive Tracker and a Meta Quest 2 controller in known positions. In this way,

<sup>2</sup>Meta Quest 2: <https://www.meta.com/it/en/quest/products/quest-2/tech-specs/>

<sup>3</sup>Unity XR Interaction Toolkit: <https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@2.4/manual/index.html>

it is possible to compare position and orientation data gathered for the two devices in the respective reference systems. Afterwards, the offsets to be applied for aligning the two reference systems are computed, and the positions and orientations of the virtual counterparts of the devices updated. In order to ensure the replicability of the calibration procedure, a physical prop (depicted in Figure 2) has been produced with 3D printing to facilitate the correct positioning of the devices.



Fig. 2. 3D printed prop used for calibrating the Meta Quest 2 and HTC Vive reference systems.

### 3.2 Virtual Curator

Particular effort has been devoted to the reconstruction of the virtual curator, which is one of the key elements of the VR experience. As illustrated in [55], a high-quality 3D model of the head of a real curator of Museo Egizio has been generated via photogrammetry. More specifically, the 3D acquisition envisaged shooting 60 photos. The 3D reconstruction and texture extraction were achieved by leveraging dedicated software such as Agisoft Metashape<sup>4</sup> and Maxon Zbrush<sup>5</sup>. Once the mesh of the head has been obtained, the full body as, well as the hair, accessories textures, and materials were added by leveraging resources in Reallusion Character Creator 3<sup>6</sup>. This software provides resources to generate realistic VHs that are compliant with the majority of the game engines. The Blender<sup>7</sup> 3D graphics suite was used to combine the different resources and prepare the avatar for animation. More specifically, the Rigify add-on<sup>8</sup> was used for rigging the avatar, i.e. connecting the 3D model to a skeleton structure made up of bones, joints, and controls (known as a rig) in order to pose and animate it. Figure 3a and Figure 3b show the generated 3D model of the virtual curator.

Once the model has been rigged, body animations were generated through the traditional keyframing technique. For the proposed experience, only two types of animations were created, i.e. idle and simple arms movements to be performed by the avatar while talking to emphasize the narration.

For what it concerns face animation, the reference work [55] leveraged motion capture. Although this technique can produce extremely realistic results, it introduces severe limitations on the flexibility/scalability of the experience. In fact, creating a new experience or even editing a single word in the narration made by the virtual curator would require the entire animation sequence to be recorded and synchronized again. For this reason, for the experience illustrated in the present paper, a different approach based on lip-sync and recorded audio

<sup>4</sup>Agisoft Metashape: <https://www.agisoft.com/>

<sup>5</sup>Maxon Zbrush: <https://www.maxon.net/en/zbrush>

<sup>6</sup>Reallusion Character Creator 3 <https://www.reallusion.com/character-creator/>

<sup>7</sup>Blender: <https://www.blender.org/>

<sup>8</sup>Rigify: <https://docs.blender.org/manual/en/3.6/addons/rigging/rigify/index.html>

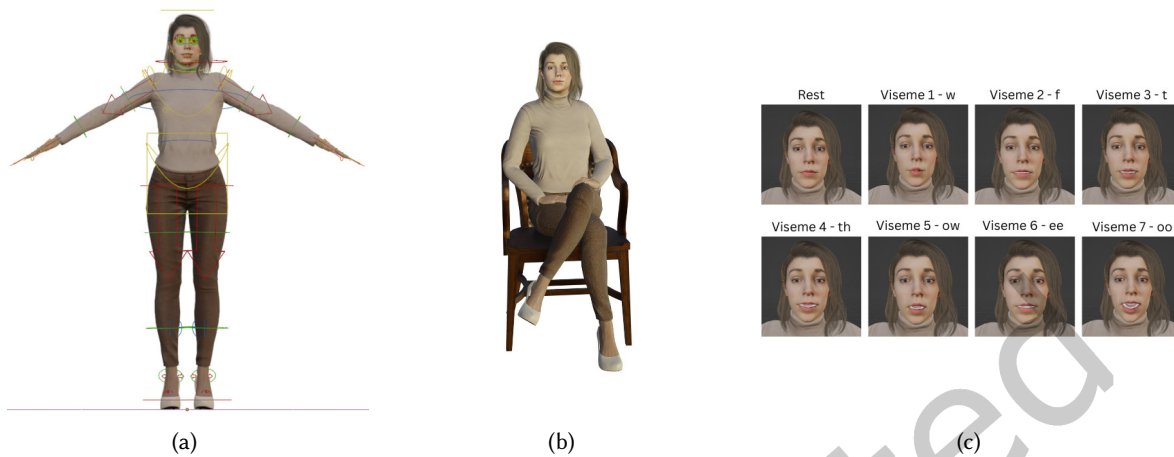


Fig. 3. 3D model of the virtual curator: a) details of the rig, b) posed avatar, and c) visemes, i.e. blendshapes, created for the virtual curator (plus rest pose).

clips was exploited. More specifically, with this approach, the movements of the avatar's lips are automatically generated in real time based on the played audio clips. In this way, it is necessary to record (and possibly re-record, to update the experience) only the audio clips. In the future, Text-to-Speech (TTS) tools could be used to speed up the process by eliminating the need to record the audio clips. In the present paper, the use of TTS tools was not considered due to the fact that they still present some limitations that may negatively impact realism like, e.g., the inadequate representation of emotions, the lack of spontaneous verbal language in terms of naturalness and comprehensibility, and the unnatural reproduction of sounds.

The lip-sync approach was implemented by leveraging the SALSA LipSync Suite for Unity<sup>9</sup>. SALSA is able to elaborate in real time an input audio file to control/animate both 2D e 3D characters. Moreover, the software offers the possibility to control the movement of the eyes, eyelids, and head to make the avatar assume secondary facial expressions that are added to the main expression to provide more variety in the face animation and make it more life-like. In order to generate the face animations, SALSA makes use of visemes, i.e., the basic units for visually representing face appearance during speech production. Visemes represent specific configurations of the mouth, lips, and tongue that occur when pronouncing certain sounds or phonemes. In this way, it is possible to use a limited number of visemes and combine them to reproduce the facial movements of an avatar during a speech. By analyzing the input audio file, SALSA recognizes the involved visemes and associates them in real time with the corresponding face shape. The sequence of viseme activations generates the resulting face animation.

The SALSA documentation suggests a standard set of visemes that can be combined to achieve a wide range of realistic expressions for lip-sync<sup>10</sup>. For the 3D model of the virtual curator, these visemes have been recreated by manipulating in Blender the complex facial rig automatically generated by the Rigify add-on, composed of 90 bones. The deformed face meshes corresponding to the visemes have been saved as blendshapes [10], which can be dynamically activated by SALSA in the Unity application to generate the lip-sync. In the case of VHs used as narrators or virtual guides, it is generally not necessary to express intense emotions through exaggerated facial expressions [70]; in fact, a pleasant experience can be achieved by ensuring that the narration is interesting and

<sup>9</sup>SALSA LipSync Suite: <https://assetstore.unity.com/packages/tools/animation/salsa-lipsync-suite-148442>

<sup>10</sup>SALSA Recommendations: <https://crazyminnowstudio.com/docs/salsa-lip-sync/modules/recommendations/>

well-structured [70]. Overall, 7 blendshapes were defined for the 3D model. Figure 3c shows the visemes and the corresponding blendshapes.

The 3D model of the virtual curator along with the animations and blendshapes were exported from Blender as a .fbx file and imported in Unity for creating the VR experience. The rig and the blendshapes are also used in Unity to trigger/perform additional animations aimed to increase the realism of the VR experience. More specifically, random micro-movements are dynamically added to make the avatar behave in a more natural (less static, or “robotic”) way. With SALSAs it was also possible to assign a target to the movement of the eyes and head, so that the virtual curator can look towards the user during the experience. Eye contact helps to increase the users’ emotional response in VR, since it gives the impression that the avatar is actually interacting with them, thus heightening the sense of presence and engagement [37].

As mentioned in Section 1, a SUI has been leveraged to let the users interact with the virtual curator. More specifically, the users can use their voice to ask questions regarding the remains. To implement the SUI, two main components were exploited.

The first component is RASA<sup>11</sup>, an open Generative Conversational Artificial Intelligence (AI) platform for creating intelligent agents. The Natural-Language Understanding (NLU) pipeline of RASA requires a text as input. The AI model handles the so-called intent classification on the text by leveraging training examples. An intent represents what the user says, whereas training data consists of examples of possible users’ utterances that are categorized by intent. Besides the default intent (that is triggered when recognition confidence is lower than an empirically defined threshold), for the proposed experience RASA was configured to identify four main intents (connected to four topics that can be presented by the virtual curator) as well as two additional intents that can be used to control the flow of the experience and receive help (more details will be provided in Section 3.5). Overall, to identify the above intents, 84 training examples were generated.

The second component is a Speech-to-Text module in charge of transforming the users’ utterances into text. The text generated from this module can be provided as input to RASA for the intent classification. To this purpose, the Microsoft Speech SDK based on Azure AI Services<sup>12</sup> was leveraged.

### 3.3 Virtual Remains

For the experience illustrated in the present paper, the objects belonging to the temporary exhibit named “Archeologia Invisibile”<sup>13</sup> hosted at the Museo Egizio from March 2019 to January 2022 were considered. This choice was made for two main reasons: research material was available to support an effective narration, and the high-quality 3D models had already been created.

In particular, among the objects in the catalogue, the mummy of a cat (Cat.2348/1) [26, 69] was selected for its size and shape, since the users could easily and comfortably manipulate it using their hands (though the experience could be easily extended to other handheld objects, not necessarily belonging to the considered exhibit or museum). Moreover, the object contains fine details on its surface, which make it extremely interesting to discover with bare hands.

The 3D model of the mummy was generated by the Museo Egizio’s researchers via photogrammetry. More specifically, the reconstruction was done by elaborating 178 photos through the 3DF Zephyr software<sup>14</sup>. Figure 4a shows a closeup of the generated 3D mesh, from which it is possible to observe the fine granularity. A solid visualization (without textures) of the model is provided in Figure 4b. For this object, both Neutron/CT scan images showing the interior of the animal remains as well as a reconstruction of the original textile were available. These resources were integrated in the VR experience, and the user was given the possibility to activate their

<sup>11</sup>RASA: <https://rasa.com/>

<sup>12</sup>Microsoft Azure Speech-to-Text: <https://learn.microsoft.com/en-us/azure/ai-services/speech-service/get-started-speech-to-text>

<sup>13</sup>Archeologia Invisibile: <https://museoegizio.it/en/explore/exhibitions/archeologia-invisibile/>

<sup>14</sup>3DF Zephyr: <https://www.3dflow.net/3df-zephyr-photogrammetry-software/>

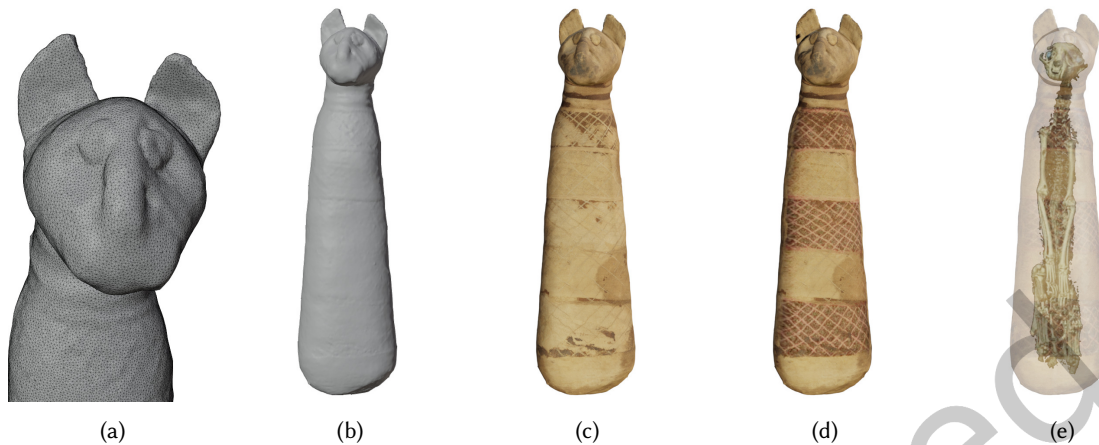


Fig. 4. Reconstruction of the mummy and related research data: a) closeup and b) and entire 3D mesh, c) current appearance, d) original texture, and e) interior with animal remains.

visualization as overlays of the virtual object. Figure 4c–4e show the different visualizations of the mummy: current appearance of the mummy as it could be seen by the visitors of the exhibit (Figure 4c), original texture that has been reconstructed through archeological studies based on information and analysis of the bandages and pigments used in Ancient Egypt (Figure 4d), and interior with the animal remains as Neutron/CT scans (Figure 4e).

As anticipated in Section 1, in the devised VR experience, PHs are used to enable a tangible interaction with the remains. The physical prop used to this purpose was produced by means of 3D printing (precisely, via Fused Filament Fabrication). As it can be observed from Figure 5a, the high-quality mesh made it possible to produce an extremely precise replica of the mummy, which can support a rich tactile experience. As said, the prop is tracked by means of an HTC Vive Tracker attached to it (Figure 5b). By leveraging the Valve Lighthouse tracking system<sup>15</sup> (consisting of two base stations), it is possible to align in real-time the physical prop and its VR counterpart, thus simultaneously providing the users with high-fidelity visual and haptic stimuli.

In the process used to convert the 3D mesh of the prop to a format compatible with the 3D printer (STL), several changes were made to the geometry in order to deal with some requirements concerning the envisaged experience. In particular, a docking system was designed to attach the HTC Vive Tracker to the prop. As indicated in the HTC Vive developer guidelines<sup>16</sup>, the Tracker can be locked to other objects and surfaces by leveraging a standard camera mount (1/4" screw nut) and the stabilizing pin recess. The designed docking system is based on a sliding tripod cradle head that can be slid into its housing on the prop and removed when needed (e.g., to move the Tracker to another prop or recharge its battery). The position for integrating the docking system in the 3D model of the mummy (on the rear bottom, as shown in Figure 5c) was carefully selected with the aim to minimize occlusions and balance the overall weight of the prop (the weight of the Tracker is 270g, though lighter and smaller alternatives are available<sup>17</sup>). A further aspect that was considered regarded the reproduction of the real artifact's weight and its distribution (about 1900). This aspect was deemed as particularly relevant to improve the tactile experience and make it more realistic. In order to make the 3D printed prop heavier (as it was

<sup>15</sup>Valve Lighthouse base stations: <https://www.valvesoftware.com/en/index/base-stations>

<sup>16</sup>HTC Vive Tracker documentation <https://developer.vive.com/resources/hardware-guides/vive-tracker-developer-guidelines/>

<sup>17</sup>Tundra Tracker: <https://tundra-labs.com/>

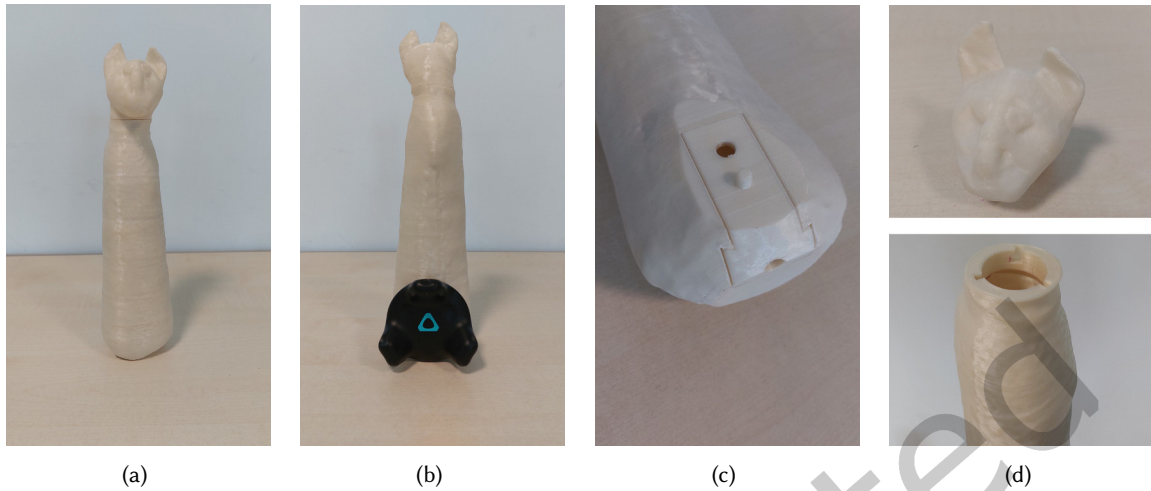


Fig. 5. PH prop used for the virtual remain of the mummy: a) replica obtained through 3D printing, b) HTC Vive Tracker for tracking it, c) docking system on the rear of the prop, and d) internal space created for filling the prop in to obtain the proper weight and locking mechanism.

<-UNO->

produced using Polylactic acid PLA), it was chosen to fill it with material(s) having the proper weight. To this aim, a “twist & lock” mechanism was designed and integrated in the model to enable several filling attempts aimed to identify the correct weight before definitively sealing the prop. Moreover, the model was modified to create a hollow space, which was filled in with a mixture of sand and styrofoam (it was ultimately decided to limit the weight to 1200g, as with the real weight users’ fatigue in the VR experience was too high). Figure 5d shows the modified model with the hollow space, and the locking mechanism.

The model was sliced in three parts (head, body and mounting system), which were printed separately using a Ultimaker S5 3D printer<sup>18</sup>. The printer was configured as follows: layer height 0.2 mm, wall thickness 1.2 mm, infill pattern zig-zag, infill density 5.0%, temperature 210°C, and speed 70 mm/s. The printing process lasted more than 10 hours.

### 3.4 Virtual Environment

For the virtual environment, the warehouses of the Museo Egizio were initially considered, but were later excluded since the location may not be capable to have a strong, visual impact on the users. Other locations in the museum would have to be visually altered to host the experience. In the end, a high-realism 3D model of an interior inspired by the Bodleian Libraries<sup>19</sup> at the Oxford University was used. This choice was meant to recall a cultural place, and was expected to be particularly captivating for the users (as confirmed in the experiments), also because it has been used as a location for some blockbuster movies. The decontextualization introduced by a virtual environment not corresponding to the place where the remain has been excavated or it is hosted today is expectfully balanced by the evocative setting that, in the collective imagination, recalls a learning experience with a magical flavour. The lighting, furniture, and books on the shelves were added to the virtual environment to prepare the user to the experience. The resulting environment is illustrated in Figure 6.

<sup>18</sup>Ultimaker S5: <https://ultimaker.com/3d-printers/s-series/ultimaker-s5/>

<sup>19</sup>Library - UE4 Project: <https://aikoshinohara.gumroad.com/l/AikoShinohara>

In order to boost the visual quality of the VR experience, Unity's High Definition Render Pipeline (HDRP) was adopted<sup>20</sup>. This choice was made to increase the degree of photorealism, since with HDRP it is possible to enhance the fidelity of lighting, shadows, and appearance of the materials, thus achieving high-quality and detailed visual rendering.



Fig. 6. Virtual environment hosting the VR experience.

### 3.5 Experience and Narration

The flow diagram in Figure 7 shows the structure of the VR experience and the narration that is currently supported.

At the beginning, a configuration procedure is envisaged in order to set up the experience and initialize the VR systems. During the configuration, which is managed through a graphics user interface (GUI) that can be operated with bare hands, it is possible to choose which virtual remains will be included in the experience. Presently, only the mummy is supported, though the application is already designed to handle other objects; in case of multiple objects, the GUI supports the user in defining the mapping with the HTC Vive Trackers attached to them. The configuration GUI is illustrated in Figure 8. The left panel, named "Remains", contains the list of virtual remains to be included in the experience. To define the above mapping, the user can choose an item from the list and associate it with one of the Trackers in the right panel named "Devices". The checkboxes on the bottom can be used to activate/deactivate the use of the TUI and the SUI (more details on these two functionalities that were added to support the experimental evaluation will be provided in Section 4).

Before starting the experience, the calibration of the VR systems has to be performed, by placing the calibration prop (with the Meta Quest 2 handheld controller and the HTC Vive Tracker in it) on the table in front of the user and pressing the Trigger button of the controller. Although it is not always required, running it right before the experience helps to reduce possible tracking drifts.

Once the configuration is complete, the experience can be started. The user finds himself or herself seated in the virtual library, in front of a wooden table (Figure 9). The virtual replica of the mummy (and other remains, if available) is placed on the table, and the virtual curator is sitting behind it. The virtual curator starts presenting herself, the experience and its content through a welcome audio clip that invites the user to take the virtual remain in his or her hands. The user is free to explore the virtual environment, interact with the virtual remain

<sup>20</sup>HDRP: <https://unity.com/srp/High-Definition-Render-Pipeline>

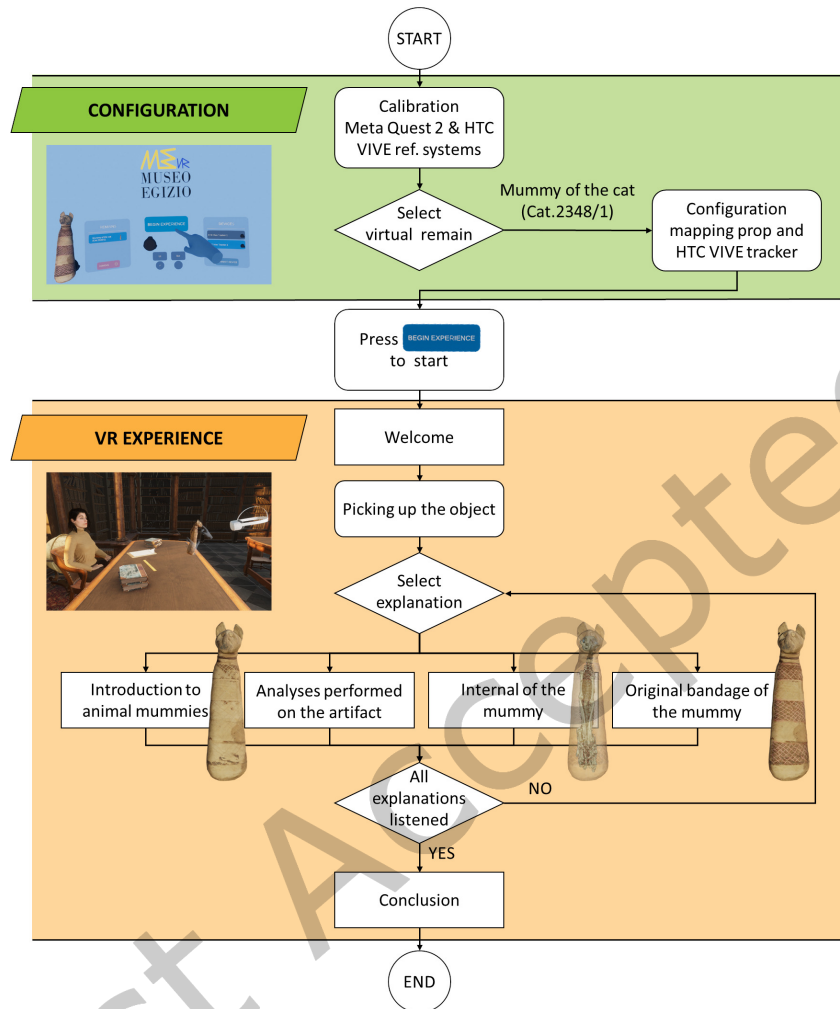


Fig. 7. Flow diagram showing the structure of the VR experience.

and discover its details. When the user picks up the object by physically grabbing its PH, the virtual curator illustrates the topics for which further explanations are available (four, at the moment), and explains how to request them using the voice. The information presented in the experience is a structured elaboration of all the material available at Museo Egizio for the selected remain. The particular remain has been chosen since some of its characteristics (e.g., the presence of elements inside the bandages and the original composition of the textile) can be experienced by the visitors using touch. The material has been organized in four main blocks with the supervision of experts of Museo Egizio, and the proposed structure allows the visitors to receive self-consistent information at once and then take a break, if needed, to elaborate the just listened content. The four topics are as follows: i) a brief introduction to animal mummification and the reasons behind this practice in Ancient Egypt; ii) the multispectral analysis and virtual bandage removal that made it possible to reconstruct the original two-tone color of the textile and safely discover what is hidden inside the mummy, respectively; iii) the studies on the



Fig. 8. GUI used for configuring the VR experience.

animal remains, which revealed the original position of the skeleton, the lack of internal organs (indicating that the animal was probably eviscerated before being mummified), and the presence of two blue-colored elements inside the orbital cavities; iv) the reconstruction of the color and other characteristics of the original bandage.

The user can ask for explanations by formulating statements or requests that contain elements recalling the above topics, e.g., “Tell me about the bandage” or “What’s inside the mummy?”. On the table, there are some books that briefly summarize each of the topics, thus helping the user to remember what the virtual curator can talk about. During the experience, the user can request the list of topics still not presented, by asking questions like “What haven’t you told me about yet?”. In order to make it easier for the user to follow the flow of the experience and provide him or her with a feedback on the percentage of completion, book titles are grayed once the associated topics have been presented (as shown in Figure 9).

The user can request explanations in any order, and can listen to them multiple times, if desired. In case the system misunderstands the user’s intention and starts presenting a wrong topic, he or she can ask the virtual curator to stop talking. Once the application recognizes that all the four topics have been presented, the virtual curator invites the user to visit the main exhibit of the museum in order to learn more about the content of the experience, and guides him or her in the removal of the headset to exit it.

#### 4 EXPERIMENTAL SETUP

In order to evaluate the proposed VR experience, a user study was carried out by involving students and staff at the authors’ university. The study was aimed to analyze the effects of leveraging the considered techniques on users’ engagement. The analysis was carried out under the hypothesis that the use of TUIs and SUIs would improve the realism and naturalness of the interaction, leading to a better overall experience.



Fig. 9. User in the proposed VR experience facing the virtual curator. Colors of the book titles indicate the topics already presented (light yellow)/still to be presented (gray).

#### 4.1 Design

In order to test the above hypothesis, the proposed experience could be compared with that in [55] (later referred to as the baseline (*BL*), which mainly differs in the way interactions are handled. In [55], only the Meta Quest 2 handheld controller were used both to manipulate the virtual objects and communicate with the avatar of the curator (as shown in Figure 10a). In order to grab an object, the user could either move a controller close to it and press the Grip button, or use the raycasting with the Trigger button to point at it and keep pressing the button to activate the grabbing; once grabbed, the object could be manipulated by simply moving the controller. To trigger the playback of the audio clips with the explanations, the buttons in a floating GUI could be activated using raycasting.

However, by directly comparing only these two modalities, it would be not possible to isolate the individual contribution brought by the considered HCI techniques, since in the proposed experience they are used together. Hence, the evaluation was designed in the form of a breakdown analysis. More specifically, starting from the proposed setup that includes both the TUI and the SUI (in the following referred to as *TUI+SUI*, two additional configurations were obtained by removing one of the interaction modalities at a time. When the TUI is removed, the users are requested to leverage the controllers for manipulating the virtual objects (as it happens in the *BL* modality), but they can use their voice for communicating with the avatar (and, like in the *BL*, books on the table recall explanations still to be presented, as shown in Figure 10b); in the following, this modality will be referred to as *onlySUI*. In turn, when the SUI is removed (*onlyTUI* modality), the users can leverage their hands and the PH to interact with the virtual objects, but the communication with the avatar is mediated by the GUI already presented for the *BL* modality; differently than in latter modality, however, the users can press the buttons in the GUI by using the hands rather than the controllers, as illustrated in Figure 10c.

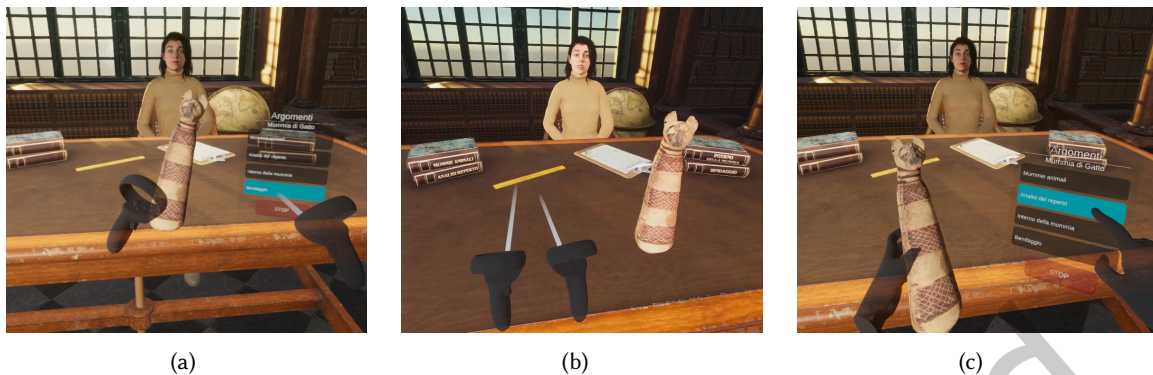


Fig. 10. Configurations compared with the proposed one: a) BL, b) onlySUI, and c) onlyTUI.

Four videos, one for each of the considered modalities (BL, onlySUI, onlyTUI, and SUI+TUI), are available for download<sup>21</sup>.

The study followed a within-subject design, hence each participant was requested to experience all the modalities. Latin square order of exposition was adopted to counterbalance potential learning effects and minimize possible biases.

#### 4.2 Participants

With the goal of determining the required sample size, an a-priori power analysis was performed using the G\*Power tool [20]. Setting  $\alpha = 0.05$  and aiming at detecting at least an effect size of medium entity (Cohen's  $f = 0.25$ ), it was found that a total sample size of 24 participants was adequate to reach a power of  $(1 - \beta) = 0.8$  for the arranged study design [18]. The 24 volunteers (16 male and 7 female) involved in the study were aged between 21 and 34 ( $M = 26.52$ ,  $SD = 3.19$ ). According to collected demographic data, 39.13% of the participants used VR devices regularly, 39.13% used them sometimes, whereas 21.74% never used them. Data also revealed that most of the participants were generally not much used to playing video games (30.43% never played them, 43.48% sometimes, 26.09% regularly) and interacting with voice assistants (34.78% never, 60.87% sometimes, 4.35% regularly). The participants were not specifically familiar with cultural heritage; in this way, it was possible to mimic the experience of common visitors of museums, who are presumed to have, on average, limited expertise in the field.

#### 4.3 Evaluation criteria

The four modalities were compared by means of subjective measurements. These measurements were collected at the end of the experiment by requesting each participant to fill in a post-test questionnaire, which is available for download<sup>22</sup>. The questionnaire included a number of sections aimed at evaluating different aspects of the experience. The first section evaluated the usability of the four modalities by means of the System Usability Scale (SUS) [7]. Questions were expressed as statements to be evaluated on a 1–5 Likert scale from “strongly disagree” to “strongly agree”. The second section was aimed to assess the participants' intention to visit the museum after the VR experience, and their willingness to participate again in the experience as part of a museum visit in the future. The statements in this section were based on the questionnaires proposed in [39] and [15], respectively,

<sup>21</sup>Videos of the proposed experience: <https://www.dropbox.com/sh/d4szuo3nqkmihwb/AABy5-tjQ9iwi3Zb7-IIA7Uza?dl=0>

<sup>22</sup>Questionnaire used for the evaluation: <https://www.dropbox.com/sh/bf8ozwa4n9ikr8g/AABHscvkv41i0UXMOL6MaisXa>

to be evaluated on a 5-point Likert scale. The third section investigated the participants' sense of immersion by means of the Immersive Experience Questionnaire (IEQ) [36]. Like in [32], a subset of 12 questions relevant for the particular type of experience, to be evaluated on a 5-point Likert scale, was used. In the fourth section, the participants' sense of presence was evaluated by means of the Presence Questionnaire (PQ) [73]. In the present study, it was chosen to focus only on the factors regarding "Involvement" and "Sensory fidelity" [71, 72], as already done in previous works (e.g., [27, 40, 43]); these items had to be evaluated on a 7-point Likert scale. The fifth section was aimed to analyze the user experience; to this aim, the 14 questions of the Game Experience Questionnaire (GEQ) (In-game module) [33] were selected. The participants were asked to indicate to what extent they felt engaged during the experience by using a 1-to-5 Likert scale from "not at all" to "extremely". Finally, the participants were requested to rank the four modalities, expressing their overall satisfaction in participating in the experiences by using each of them.

#### 4.4 Procedure

Upon arrival, the participants were informed about the experimental procedure. After collecting the informed consent and demographic data, they were provided with brief instructions on how to use the VR application in the four modalities. The participants were then requested to wear the VR headset and begin the experience. During the experience, no restrictions were set regarding the activation of the audio explanations for the four topics (Section 3.5); thus, they could trigger the same explanation several times, in any order, and could stop the playback at any time. When they had listened to all the explanations, the experience was considered as completed. At the end of the experience, the participants were asked to fill in the post-test questionnaire and then move to the other modalities. At the conclusion of the experiment, they were asked to express their ranking for the four modalities.

## 5 RESULTS

The statistical significance of the obtained results was analyzed by performing the Friedman test ( $p$ -value < 0.05) with the Wilcoxon signed-rank test for paired samples as post-hoc. The effect size was measured through Cohen's  $d$ .

As said, the first section of the questionnaire requested the participants to evaluate the usability of the four modalities using the SUS [7]. Based on collected results (BL: 74.38, onlyTUI: 83.13, onlySUI: 78.02, SUI+TUI: 91.46;  $p < .001$ ), the participants found the SUI+TUI modality more usable than the BL ( $p < .001$ ,  $d = -1.329$ ), onlyTUI ( $p = .014$ ,  $d = -0.658$ ), and onlySUI ( $p < .001$ ,  $d = -1.162$ ) ones. Moreover, the onlyTUI modality was perceived as more usable than the BL one ( $p = .044$ ,  $d = -0.526$ ). The scores assigned to each statement are reported in Figure 11. According to the categorization proposed in [3], the scores obtained by the four modalities correspond to the following grades (adjective rating): BL: B (Good), onlyTUI: A (Excellent), onlySUI: B+ (Good), SUI+TUI: A+ (Excellent).

Regarding the second section of the questionnaire based on [15, 39], an overall score was obtained for the intention to visit the museum and intention to use again the VR experience in future visits of the museum by averaging the scores assigned to each statement (it is worth noticing that, to compute average values and for the sake of readability, the scores for statements in a negative form in the second and in the remaining sections have been reversed, mapping all the values on a worse-to-better scale). Overall, significant differences were observed for both the intention to visit (BL: 3.18, onlyTUI: 3.53, onlySUI: 3.23, SUI+TUI: 3.72;  $p < .001$ ) and the intention to use again the VR experience in future visits (BL: 3.76, onlyTUI: 4.08, onlySUI: 3.82, SUI+TUI: 4.25;  $p < .001$ ). More specifically, it was observed that the participants were more willing to visit the museum after having used the SUI+TUI modality than the BL ( $p = .001$ ,  $d = -0.627$ ), onlyTUI ( $p = .022$ ,  $d = -0.228$ ), and onlySUI ( $p < .001$ ,  $d = -0.570$ ) ones. Moreover, the onlyTUI modality was preferred to the onlySUI ( $p = .011$ ,  $d = -0.379$ )

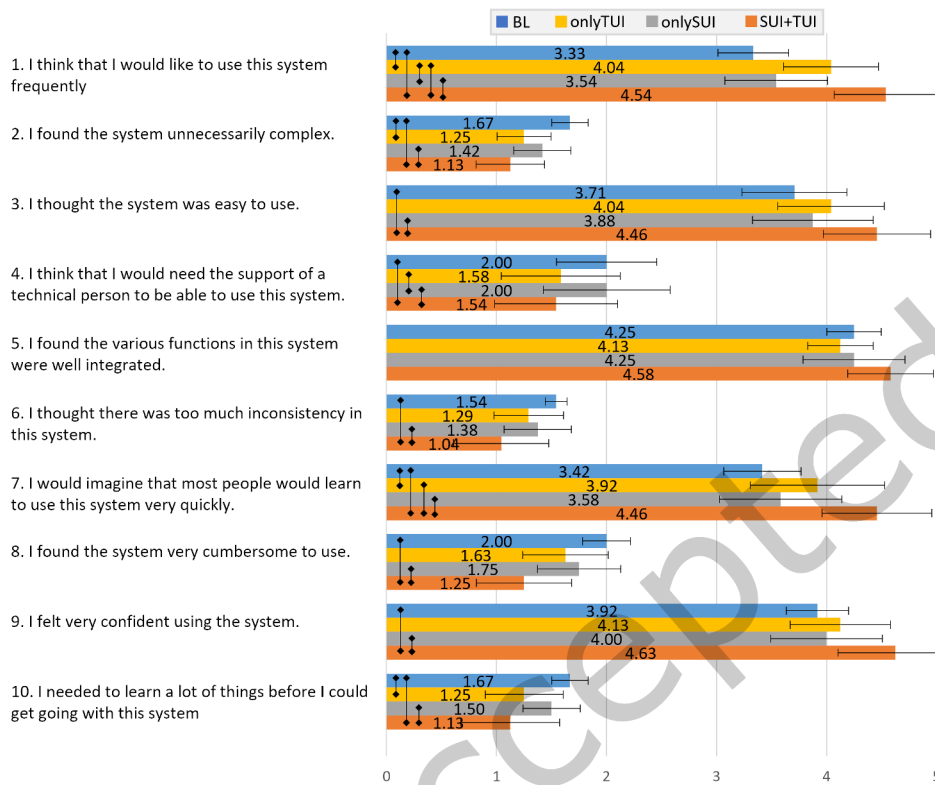


Fig. 11. Subjective results in terms of usability based on the SUS [7]

and BL ( $p = .004$ ,  $d = -0.442$ ) ones. Concerning the intention to use again the VR experience in future visits, the participants preferred more the SUI+TUI modality than the BL ( $p = .002$ ,  $d = -0.614$ ), onlyTUI ( $p = .034$ ,  $d = -0.225$ ), and onlySUI ( $p = .004$ ,  $d = -0.543$ ) ones. Moreover, post-hoc significant differences were observed between the onlyTUI modality and both the BL ( $p = .006$ ,  $d = -0.399$ ) and onlySUI ( $p = .013$ ,  $d = -0.329$ ) ones. For both the dimensions in this section, no significant differences were observed between the BL and onlySUI modalities.

Results regarding the sense of immersion based on the IEQ [36] are reported in Figure 12. Overall, significant differences were found (BL: 3.17, onlyTUI: 3.42, onlySUI: 3.33, SUI+TUI: 3.72;  $p < .001$ ). In particular, the participants reported that when using the SUI+TUI modality they had the perception to be more immersed in the VR experience than with the BL ( $p < .001$ ,  $d = -1.342$ ), onlyTUI ( $p = .001$ ,  $d = -0.654$ ), and onlySUI ( $p < .001$ ,  $d = -0.905$ ) ones. Moreover, the BL modality was found to stimulate a lower sense of immersion than the onlyTUI ( $p = .001$ ,  $d = -0.618$ ) and onlySUI ( $p = .002$ ,  $d = -0.444$ ) ones.

With respect to the sense of presence, evaluated through the PQ [73], statistically significant differences were observed for both the considered factors, i.e. involvement (BL: 4.71, onlyTUI: 5.52, onlySUI: 5.15, SUI+TUI: 6.18;  $p < .001$ ) and sensory fidelity (BL: 3.64, onlyTUI: 6.28, onlySUI: 3.72, SUI+TUI: 6.32;  $p < .001$ ). Results are detailed in Figure 13a and Figure 13b. Starting from the factor related to the involvement, it was observed that all the differences analyzed through a post-hoc analysis were statistically significant. In particular, the participants judged the SUI+TUI modality as able to make them more involved in the VR experience than the BL ( $p < .001$ ,

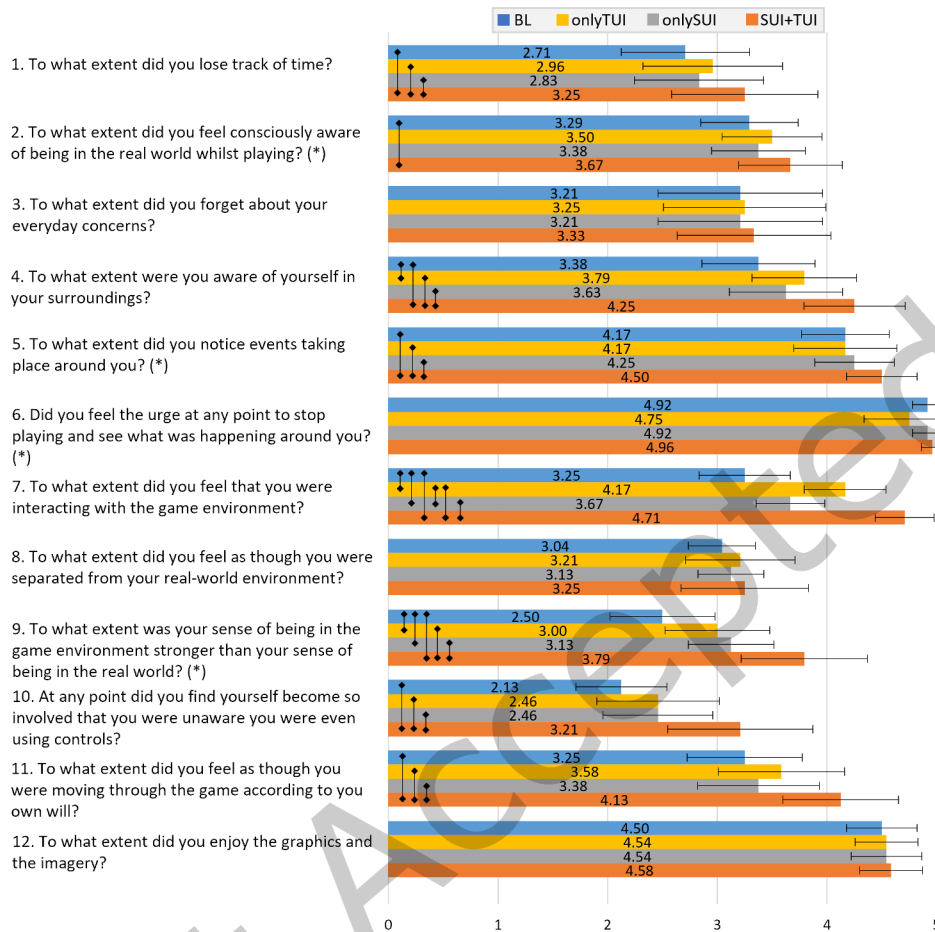


Fig. 12. Subjective results in terms of sense of immersion based on the IEQ [36]

$d = -2.232$ ), onlyTUI ( $p < .001$ ,  $d = -0.815$ ), and onlySUI ( $p < .001$ ,  $d = -1.545$ ) ones. Furthermore, the onlyTUI modality was preferred to both the BL ( $p = .001$ ,  $d = -0.987$ ) and onlySUI ( $p = .007$ ,  $d = -0.452$ ) ones, whereas the onlySUI modality made the participants feel more involved than the BL one ( $p < .001$ ,  $d = -0.636$ ). Moving to the sensory fidelity factor, the SUI+TUI modality was perceived as capable of stimulating the senses more faithfully than the BL ( $p < .001$ ,  $d = -3.404$ ) and onlySUI ( $p < .001$ ,  $d = -3.452$ ) ones. Moreover, the onlyTUI modality was judged to be characterized by a sensory fidelity higher than the BL ( $p < .001$ ,  $d = -3.402$ ) and onlySUI ( $p < .001$ ,  $d = -3.454$ ) ones. No statically significant differences were observed between the SUI+TUI and onlyTUI modalities, as well as between the BL and onlySUI ones.

For what it concerns the overall user experience (Figure 14), evaluated through the statements of the GEQ [36], statistically significant differences were observed (BL: 4.10, onlyTUI: 4.23, onlySUI: 4.25, SUI+TUI: 4.53;  $p < .001$ ). More specifically, the participants reported a better user experience with the SUI+TUI modality than the BL ( $p < .001$ ,  $d = -1.274$ ), onlyTUI ( $p = .002$ ,  $d = -0.707$ ), and onlySUI ( $p < .001$ ,  $d = -0.828$ ) ones. The post-hoc analysis also indicated significant differences between the BL and onlySUI modalities ( $p = .012$ ,  $d = -0.380$ ).

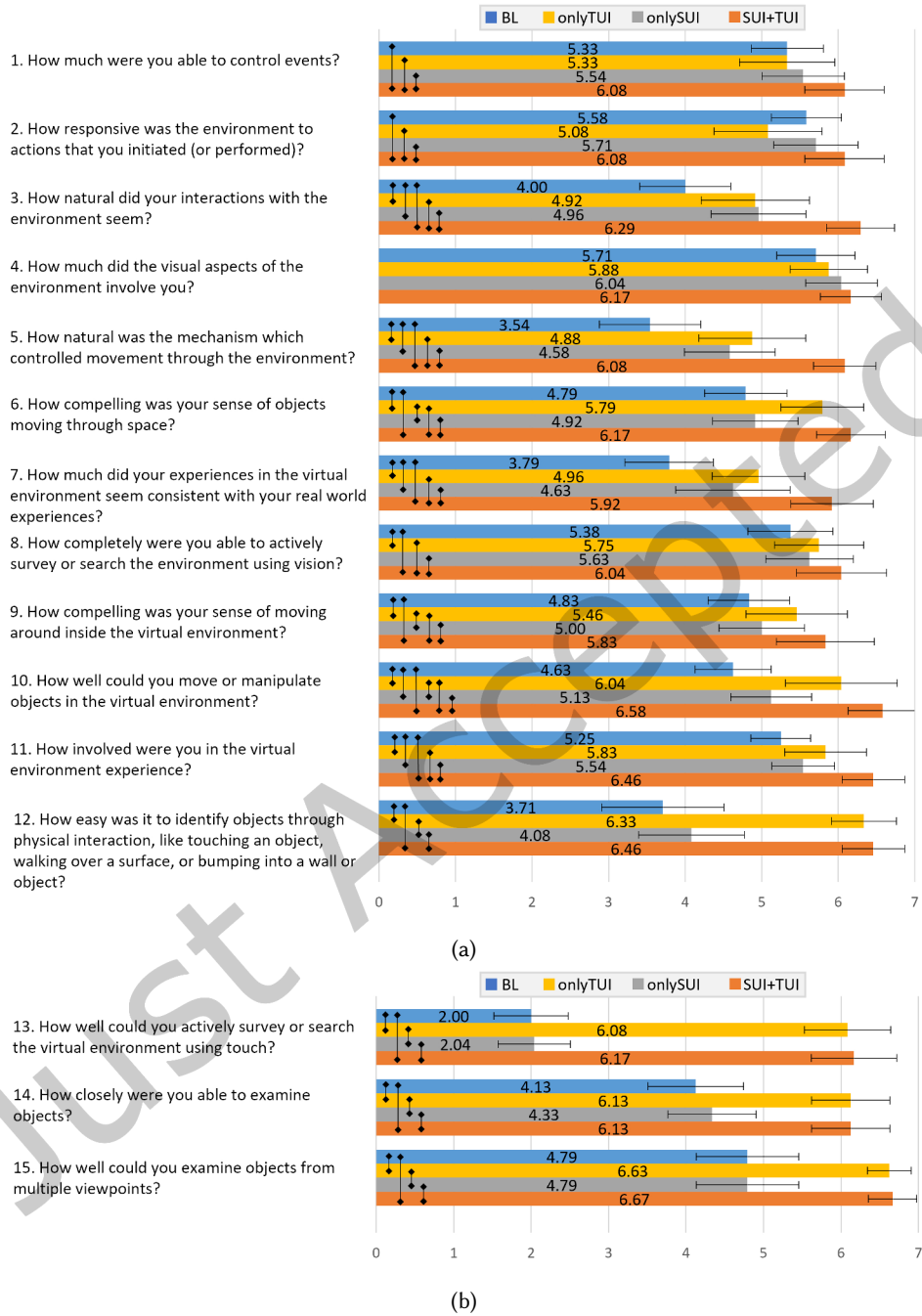


Fig. 13. Subjective results in terms of a) user involvement and b) sensory fidelity based on PQ [73].

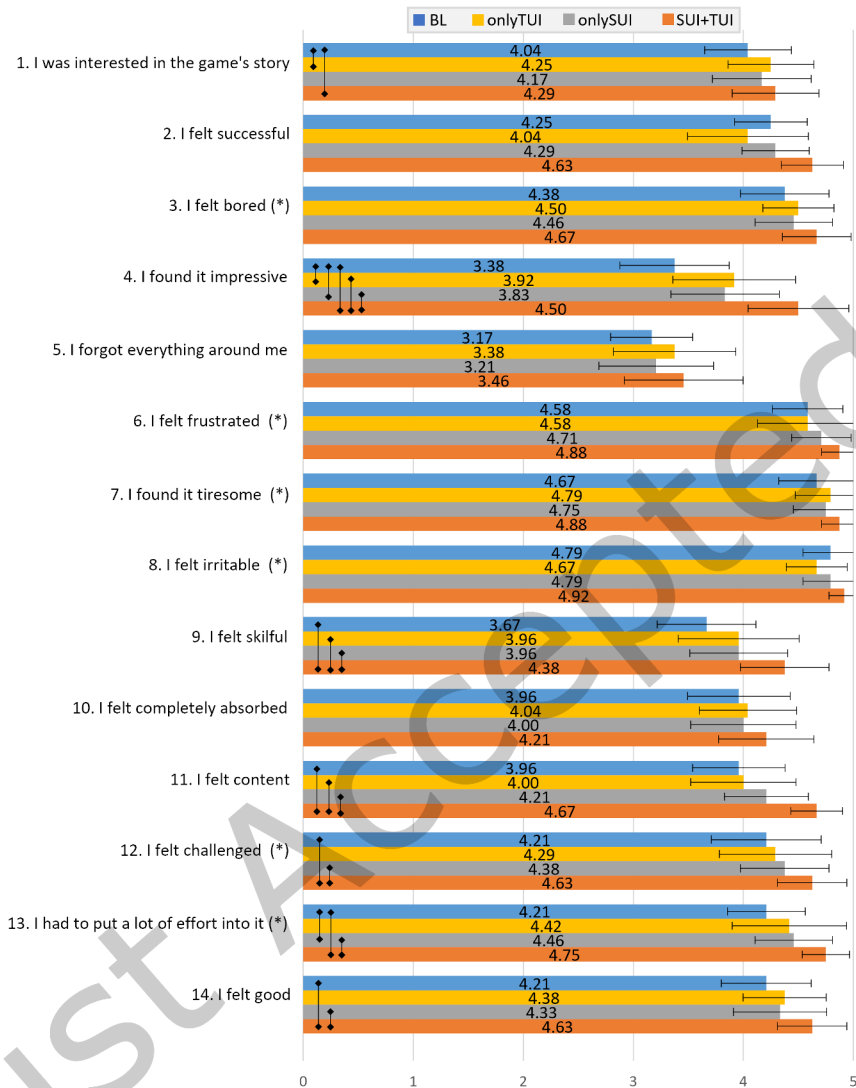


Fig. 14. Subjective results in terms of user experience based on the In-game module of the GEQ [33]

Finally, the distribution of the preferences reported by the participants at the end of the experiment are shown in Figure 15. The statistical analysis produced the following ranking: 1. SUI+TUI, 2/3. onlyTUI and onlySUI (tie) 4. BL ( $p < .001$ ).

## 5.1 Discussion

Based on the summary in the previous sub-section, it can be stated that the SUI+TUI modality was judged as superior compared to the other modalities for most of the studied dimensions. In order to dig into the motivations

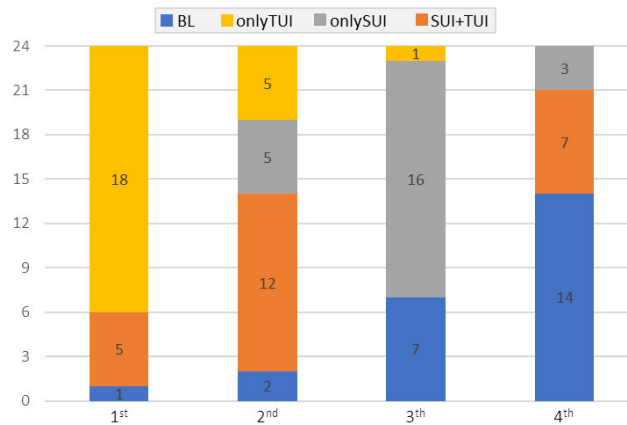


Fig. 15. Distribution of the participants' responses regarding the overall ranking of the four modalities.

behind the high appreciation for this modality, it is possible to look at the scores assigned by the participants to individual items of the various questionnaire sections.

Starting from the statements regarding the SUS (Figure 11), it can be noticed that, generally, the usability of the SUI+TUI modality was rated higher than that of the other modalities, especially the BL one. A reasonable motivation could be related to the ease of learning and use of this modality with respect to other ones. In fact, the comparison of BL and SUI+TUI modalities reveals that the latter was judged as characterized by a lower complexity (Q2,  $p = .001$ ), as easier to use (Q3,  $p = .005$ ) and learn (Q7,  $p < .001$ ; Q10,  $p < .001$ ), as requesting lower help to be used (Q4,  $p = .006$ ), and as showing a less cumbersome interaction (Q8,  $p = .001$ ). The improved interaction enabled by the SUI+TUI modality made the participants feel more confident using it (Q9,  $p = .013$ ) and judge it as characterized by a lower inconsistency (Q6,  $p = .038$ ) than the BL one. Interestingly, all these differences were also observed comparing the SUI+TUI modality with the onlySUI one; this finding may indicate that the difficulties in using the BL and onlySUI modalities could mainly derive from the use of the controllers. These results could be partly due to the relatively low experience of the participants with VR systems, since less than 40% of them stated to use this technology regularly. In particular, as indicated in the comments provided at the end of the experiments, the participants who were not familiar with VR lamented an increased cognitive demand associated with the need to remember the mapping between the application functionalities and the controllers' buttons; for this reason, when using the BL and onlySUI modalities, the participants needed more time to familiarize with the interface than when using the SUI+TUI modality.

The high usability of the SUI+TUI made the participants more willing to use the experience frequently with this modality than with the other ones, as indicated by the scores assigned to item Q1 of the SUS ( $p < .001$ ) as well as to items in the second section of the questionnaire concerning the intention to use. Moreover, the results of the second section show that the SUI+TUI modality increased the interest of the participants in visiting the museum after the VR experience more than the other modalities. This result is probably linked to the difficulties that the participants faced in interacting with the application when using, in particular, the BL and onlySUI modalities. The lower effort required by the SUI+TUI and onlyTUI modalities allowed the participants to focus more on the content of the experience, thus stimulating their curiosity to explore the rest of the collection.

Looking closer at the individual results regarding sense of immersion, it is possible to notice that the scores of individual items assigned to the SUI+TUI modality were generally higher than those of the other modalities. Figure 12 shows a clear preference for the SUI+TUI modality with respect to the other ones, e.g., for items Q1

( $p < .001$ ), Q5 ( $p = .009$ ), Q10 ( $p < .001$ ), and Q11 ( $p < .001$ ). These items highlight its ability to decouple the real and virtual environments (Q1 and Q5) by immersing the participants in a seamless virtual experience in which they were allowed to use their own voice and grab the objects with their hands as they would do in the real world (Q10 and Q11). The analysis of items Q4 ( $p < .001$ ), Q7 ( $p < .001$ ), and Q9 ( $p < .001$ ), which regard the sense of embodiment in the VR experience, shows a clear dominance of the SUI+TUI modality on the other ones. Considering items Q4, Q7, and Q9 it is also possible to notice that the introduction of even just the TUI or the SUI led to an increased sense of immersion, as statistically significant differences were also observed considering the BL, onlyTUI and onlySUI modalities.

Moving to the fourth section of the questionnaire, results indicate that the introduction of natural interaction mechanisms enhanced the sense of presence. Analyzing the individual items, it emerges that this outcome can be related to a number of factors. First, as already discussed when analyzing usability, the participants reported to be more confident in using the SUI+TUI than the other modalities, as they found it to offer more control and awareness of the virtual environment. This aspect is also confirmed by the scores assigned to Q1 ( $p = .007$ ), Q2 ( $p = .002$ ), and Q8 ( $p < .001$ ) of the PQ (Figure. 13a). The strength of the SUI+TUI modality mainly derive from the ability of this modality to enable more natural interactions with the virtual environment (Q3,  $p < .001$ ) and improve the mechanisms for controlling the virtual object (Q5,  $p < .001$ ). As a result, the operations made with the SUI+TUI modality appeared to be more consistent with the real world ones (Q7,  $p < .001$ ), and the participants felt to be more involved in the virtual experience (Q11,  $p < .001$ ). Considering items Q3, Q5, Q7 and Q11, it can be observed that the BL modality was considered as the worst one for what it concerns this aspect, as natural interactions are missing. Moreover, scores assigned to item Q10 not only confirm the higher preference of the participants for the SUI+TUI modality over the BL one ( $p < .001$ ) but, since statistically significant differences were also observed considering all the modalities, they also indicate that the lack of physical interactions in the onlySUI modality affected the sense of presence, leading the participants to prefer the onlyTUI one.

Considering the items in this section that asked the participants to rate how compelling were some aspects of the VR experience, the SUI+TUI modality was preferred to all the other modalities both for controlling the virtual object (Q6,  $p < .001$ ) and to move in the environment (Q9,  $p < .001$ ); it is worth observing that, since the VR experience expects the users to remain seated, for the sake of the performed evaluation the latter item was associated with the operations performed by the participants to ask the virtual curator for explanations, as well as explore the environment using the head/gaze. Results above can be explained by the mentioned usability issues and by the need to interact with the GUI, which forced the participants to perform movements to point at (with the BL modality) or press (with the onlyTUI modality) the virtual buttons. Moreover, the GUI could partially occlude the visualization of the virtual environment, thus requesting the participants to move it away to clear their view. Statistically significant differences were also observed between the onlySUI and onlyTUI modalities, with high preference expressed for the latter. This outcome suggests that the lack of haptic feedback reduced the interest of the participants in interacting with the virtual object and moving in the environment, since the use of the PH enabled a more compelling interaction than the controllers. It was not surprising to find statistically significant differences for the item regarding physical interaction (Q12,  $p < .001$ ), since the use of the PH helped the participants to identify the characteristics of the remain (i.e. shape and surface features) with their hands. This is reasonably the motivation that made the participants rate better the onlyTUI and SUI+TUI modalities than the BL and onlySUI ones.

Finally, the possibility to touch the remains with own hands also improved the sensory fidelity of the modalities characterized by the use of PHs (SUI+TUI and only PH) compared to the modalities based on the controllers (BL and onlySUI). As shown in Figure 13b, the participants found it easier to survey the environment using touch (Q1,  $p < .001$ ), examine the virtual object (Q2,  $p < .001$ ), and observe it from different viewpoints (Q3,  $p < .001$ ).

Considering the user experience, the last section of the questionnaire shows that, in general, the participants appreciated the VR experience, as they assigned scores higher than four to all the modalities. Focusing on

statistically significant differences, it emerges that the SUI+TUI and onlyTUI modalities made the participants more interested in the content of the experience (Q1,  $p = .038$ ). The BL modality was considered a quite common way to interact with a software application, whereas combining the TUI and SUI in the same application was recognized to provide an experience more impressive than that offered by the other modalities (Q4,  $p < .001$ ). The ease of use of the SUI+TUI modality, which does not request to use the controllers or to interact with a GUI, is highlighted by items Q9 ( $p = .011$ ), Q12 ( $p = .029$ ), and Q13 ( $p = .007$ ). At the end of the VR experience, the participants expressed a higher level of satisfaction for this modality, as confirmed by statements Q11 ( $p < .001$ ) and Q14 ( $p = .008$ ).

Finally, the overall ranking confirmed the general trend observed in the previous sections. More specifically, the majority of the participants (75%) rated the SUI+TUI as the most preferred modality. Although the onlyTUI modality was rated as the second choice by the majority of the participants (50%) – compared to the 20.83% of the onlySUI modality – no statistically significant differences were found between these two modalities; 58.33% of the participants considered the BL as the worst modality.

## 6 CONCLUSIONS AND FUTURE WORK

The present paper illustrates the design of a high-fidelity VR experience in which the users are accompanied in the discovery of remains belonging to the collection of Museo Egizio in Turin by a virtual curator. The paper extends the work in [55], where a preliminary implementation of the experience was presented in which interaction was based solely on the VR handheld controllers and the curator presented the content associated with the objects in a pre-defined order.

In the present paper, the focus is on boosting the sense of immersion and presence in the VR experience by leveraging natural HCI techniques. In particular, the use of TUIs and of SUIs is investigated. TUIs, in the form of PH props, are exploited to allow the users to explore the remains not only visually, but also physically by manipulating their 3D-printed replicas and feeling their shape, surface, size, etc.; SUIs, in turn, are leveraged to let the users communicate with the curator using their voice and ask for explanations about the remains in the preferred order. Although the implementation considered Ancient Egypt remains, the architecture supporting it is general, and could be easily applied to different handheld objects of other museums. It is worth noticing that results have been achieved by considering participants who represented a specific age range and were generally not familiar with cultural heritage. However, the literature shows that results related to the application of digital technologies could be age-dependent [14] and influenced by the previous knowledge of the participants of the specific domain [48]. Hence, in the future, the user study could be extended to take into account users of different ages and backgrounds, thus resembling a wider target of museum visitors.

A user study has been carried out as a breakdown analysis to assess the impact brought by the incremental introduction of the considered interaction techniques on users' engagement. Results showed that the use of TUIs and SUIs can significantly increase usability, as well as perceived sense of immersion, presence, and user experience compared to controller-based interaction, also raising the users' interest in visiting the museum or using this kind of experiences in future visits. The study also revealed that the introduction of only one of the analyzed interfaces can have a positive effect on engagement; although a clear advantage of one of the two techniques on the other could not be found (as indicated also by the final ranking), it was noticed that the contribution brought by the TUI was more important than that of the SUI, especially at improving the sense of presence.

Future works will be devoted to evaluating alternative interaction paradigms that may be used when the noise of the real environment prevents the correct recognition of the users' speech. Possible alternatives could consider the use of eye gaze, as proposed in other VR experiences like, e.g., Sky VR: Hold the World<sup>23</sup>; this way, social

<sup>23</sup>Sky VR: Hold the World: <https://www.oculus.com/experiences/rift/2331434793563555/>

interaction with the virtual curator could also be boosted. Moreover, the interaction with the GUI in the modality leveraging only the TUI (which was found to reduce immersion since it does not provide haptic feedback to button presses) could be improved by building the interface on the physical prop (like in [32]), or moving the virtual buttons close to the surface of the remain (so that the users can press the button and feel the surface of the passive haptic). Finally, considering the storytelling, alternative experiments will be carried out to determine whether it is preferable for the users to choose the topic they are interested into (like in the current implementation) or to have a pre-defined sequence of explanations (which would make the experience less interactive, thus possibly less engaging): in the former case, it will be necessary to evaluate approaches capable of helping the users to identify the topics that are still to be presented (the approach currently adopted based on book titles can be effective to avoid missing content in the experience, but would hardly scale with a larger number of explanations per object).

## ACKNOWLEDGMENTS

The authors want to thank Simone Restivo (Politecnico di Torino, Dipartimento di Automatica e Informatica, Italy) and Davide Mezzino (Università Telematica Internazionale Uninettuno, Facoltà di Beni Culturali, Italy) for their contribution to the design and implementation of the software in [55], used as baseline (BL) modality in the present work. Moreover, the authors want to provide credit to Simone Restivo, who generated the high-quality 3D model of the head of the curator and Aiko Shinohara (Senior Environment Palette Artist at Firewall Studios) who built the virtual environment used in this work. For what it concerns the virtual remain, the considered mummy was studied by a multidisciplinary team composed of the following researchers: Salima Ikram (AUC) and Alberto Valazza (University of Turin) for the zoological part; Matilde Borla (Superintendency of Archaeology, Fine Arts and Landscape for the metropolitan city of Turin) and Cinzia Oliva (Restoration of ancient fabrics); Debora Angelici (independent researcher) for the archaeometric analyses within the context of the project “Turin Animal Mummies” supervised by Salima Ikram, Sara Aicardi and Federica Facchetti. Benjamin Moreno accomplished the virtual unwrapping of the bandages, whereas the 3D model was created by Riccardo Antonino (Creative Art Director, Robin Studio).

The authors are also grateful to Francesca Ronco (Universitat Politècnica de València, Spain) for her support with the 3D printing of the PH.

This research was supported by the VR@POLITO lab of Politecnico di Torino, Dipartimento di Automatica e Informatica, Italy (scientific director Prof. Fabrizio Lamberti), which contributed to the design and implementation of the software. Moreover, the MOD Lab Arch of Politecnico di Torino, Dipartimento di Architettura e Design, Italy (scientific director Prof. Marco Vitali) contributed to the 3D printing of the PH. Alberto Cannavò received funding from PON “Ricerca e Innovazione” 2014-2020 – DM 1062/2021.

## REFERENCES

- [1] Alonzo C Addison. 2000. Emerging trends in virtual heritage. *IEEE Multimedia* 7, 2 (2000), 22–25.
- [2] Hamza A Al-Jundi, and Emad Y Tanbour. 2022. A framework for fidelity evaluation of immersive virtual reality systems. *Virtual Reality* 26, 3 (2022), 1103–1122.
- [3] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies* 4, 3 (2009), 114–123.
- [4] Mafkereseb Kassahun Bekele, and Erik Champion. 2019. A comparison of immersive realities and interaction methods: Cultural learning in virtual heritage. *Frontiers in Robotics and AI* 6 (2019), 91.
- [5] Andrea Bönsch, David Hashem, Jonathan Ehret, and Torsten W Kuhlen. 2021. Being guided or having exploratory freedom: User preferences of a virtual agent’s behavior in a museum. In *Proc. 21st ACM International Conference on Intelligent Virtual Agents*. 33–40.
- [6] Davide Borra, Maurizio Forte, Eva Pietroni, and Claudio Rufa. 2002. The virtual reality project of the Scrovegni Chapel of Padua. In *Proc. 8th International Conference on Virtual Systems and Multimedia*. 25–27.
- [7] John Brooke. 1996. SUS: A “quick and dirty” usability scale. *Usability Evaluation in Industry* 189, 3 (1996), 189–194.
- [8] Nikos Bubaris. 2014. Sound in museums—museums in sound. *Museum Management and Curatorship* 29, 4 (2014), 391–402.

- [9] Davide Calandra, F Gabriele Praticcò, Alberto Cannavò, Luca Micelli, and Fabrizio Lamberti. 2019. Building reconfigurable passive haptic interfaces on demand using off-the-shelf construction bricks. In *Proc. IEEE Conference on Virtual Reality and 3D User Interfaces*. IEEE, 1403–1404.
- [10] Alberto Cannavò, Emanuele Stellini, Congyi Zhang, and Fabrizio Lamberti. 2023. A sketch-based interface for facial animation in immersive virtual reality. *International Journal of Human-Computer Interaction* (2023), 1–19.
- [11] Marcello Carrozzino, Marianna Colombo, Franco Tecchia, Chiara Evangelista, and Massimo Bergamasco. 2018. Comparing different storytelling approaches for virtual guides in digital immersive museums. In *Proc. 5th International Conference on Augmented Reality, Virtual Reality, and Computer Graphics*. 292–302.
- [12] Marcello Carrozzino, Cristian Lorenzini, Mihai Duguleana, Chiara Evangelista, Raffaello Brondi, Franco Tecchia, and Massimo Bergamasco. 2016. An immersive VR experience to learn the craft of printmaking. In *Proc. 3rd Int. Conf. on Augmented Reality, Virtual Reality, and Computer Graphics*. 378–389.
- [13] Thomas Chatzidimitris, Damianos Gavalas, and Despina Michael. 2016. SoundPacman: Audio augmented reality in location-based games. In *Proc. 18th Mediterranean Electrotechnical Conference*. 1–6.
- [14] Stephanie Hui-Wen Chuah. 2018. Why and who will adopt extended reality technology? Literature review, synthesis, and future research agenda. *Literature Review, Synthesis, and Future Research Agenda (December 13, 2018)* (2018).
- [15] Namho Chung, Heejeong Han, and Youhee Joun. 2015. Tourists’ intention to visit a destination: The role of augmented reality (AR) application for a heritage site. *Computers in Human Behavior* 50 (2015), 588–599.
- [16] JE Cichor, M Egorov, DA Plecher, E Schmid, and C Peus. 2019. Everything Starts with a Handshake: Effects of Character Design and Character Interactions on Leadership Development in Virtual Reality. In *5th International Augmented Reality & Virtual Reality Conference (IAVR)*.
- [17] Leon Ciechanowski, Aleksandra Przegalinska, Mikolaj Magnuski, and Peter Gloor. 2019. In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems* 92 (2019), 539–548.
- [18] Jacob Cohen. 2013. *Statistical power analysis for the behavioral sciences*. Academic Press.
- [19] Kira Eghbal-Azar, Martin Merkt, Julia Bahnmueller, and Stephan Schwan. 2016. Use of digital guides in museum galleries: Determinants of information selection. *Computers in Human Behavior* 57 (2016), 133–142.
- [20] Franz Faul, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. 2009. Statistical power analyses using G\* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods* 41, 4 (2009), 1149–1160.
- [21] Martin Feick, Scott Bateman, Anthony Tang, André Miede, and Nicolai Marquardt. 2020. Tangi: Tangible proxies for embodied object exploration and manipulation in virtual reality. In *Proc. IEEE International Symposium on Mixed and Augmented Reality*. 195–206.
- [22] Andrea Ferracani, Marco Faustino, Gabriele Xavier Giannini, Lea Landucci, and Alberto Del Bimbo. 2017. Natural experiences in museums through virtual reality and voice commands. In *Proc. 25th ACM international conference on Multimedia*. 1233–1234.
- [23] Carola Gatto, Sofia Chiarello, Federica Faggiano, Benito Luigi Nuzzo, Ileana Riera Panaro, Giada Sumerano, and Lucio Tommaso De Paolis. 2023. Enhancing Accessibility of Cultural Heritage: Extended Reality and Tactile Prints for an Inclusive Experience of the Madonna Dell’Itri Church in Nociglia. In *International Conference on Extended Reality*. Springer, 146–159.
- [24] Joe Geigel, Kunal Shailesh Shitut, Juilee Decker, Amanda Doherty, and Gary Jacobs. 2020. The digital docent: XR storytelling for a living history museum. In *Proc. 26th ACM Symposium on Virtual Reality Software and Technology*. 1–3.
- [25] Jonathan Gratch, Ning Wang, Anna Okhmatovskaia, Francois Lamothe, Mathieu Morales, Rick J van der Werf, and Louis-Philippe Morency. 2007. Can virtual humans be more engaging than real ones?. In *Proc. 12th International Conference on Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments*. 286–297.
- [26] Christian Greco. 2019. Keynote: The biography of objects. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 42 (2019), 5–10.
- [27] MP Jacob Habgood, David Moore, David Wilson, and Sergio Alapont. 2018. Rapid, continuous movement between nodes as an accessible virtual reality locomotion technique. In *Proc. IEEE conference on Virtual Reality and 3D User Interfaces*. 371–378.
- [28] Ramy Hammady, Minhua Ma, Carl Strathern, and Mostafa Mohamad. 2020. Design and development of a spatial mixed reality touring guide to the Egyptian museum. *Multimedia Tools and Applications* 79 (2020), 3465–3494.
- [29] Hunter G Hoffman. 1998. Physically touching virtual objects using tactile augmentation enhances the realism of virtual environments. In *Proc. IEEE Virtual Reality Annual International Symposium (Cat. No. 98CB36180)*. 59–63.
- [30] Eva Hornecker. 2011. The role of physicality in tangible and embodied interactions. *Interactions* 18, 2 (2011), 19–23.
- [31] David Howes. 2014. Introduction to sensory museology. *The Senses and Society* 9, 3 (2014), 259–267.
- [32] Vedad Hulusic, Linda Gusia, Nita Luci, and Michael Smith. 2022. Tangible user interfaces for enhancing user experience of virtual reality cultural heritage applications for utilization in educational environment. *ACM Journal on Computing and Cultural Heritage* 16, 2 (2022).
- [33] Wijnand A IJsselsteijn, Yvonne AW De Kort, and Karolien Poels. 2013. The game experience questionnaire. (2013).
- [34] Chiara Innocente, Luca Ulrich, Sandro Moos, and Enrico Vezzetti. 2023. A framework study on the use of immersive XR technologies in the cultural heritage domain. *Journal of Cultural Heritage* 62 (2023), 268–283.

- [35] Hiroshi Ishii and Brygg Ullmer. 1997. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems*. ACM, 234–241.
- [36] Charlene Jennett, Anna L Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. 2008. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies* 66, 9 (2008), 641–661.
- [37] Daphne E Karreman, Elisabeth MAG van Dijk, and Vanessa Evers. 2012. Using the visitor experiences for mapping the possibilities of implementing a robotic guide in outdoor sites. In *Proc. 21st IEEE International Symposium on Robot and Human Interactive Communication*. 1059–1065.
- [38] Effie Karuzaki, Nikolaos Partarakis, Nikolaos Patsiouras, Emmanouil Zidianakis, et al. 2021. Realistic virtual humans for cultural heritage applications. *Heritage* 4, 4 (2021), 4148–4171.
- [39] Myung Ja Kim, Choong-Ki Lee, and Timothy Jung. 2020. Exploring consumer behavior in virtual reality tourism using an extended stimulus-organism-response model. *Journal of Travel Research* 59, 1 (2020), 69–89.
- [40] Myungho Lee, Gerd Bruder, and Gregory F Welch. 2017. Exploring the effect of vibrotactile feedback through the floor on social presence in an immersive virtual environment. In *Proc. IEEE Conference on Virtual Reality and 3D User Interfaces*. IEEE, 105–111.
- [41] Karl F MacDorman and Debaleena Chattopadhyay. 2016. Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. *Cognition* 146 (2016), 190–205.
- [42] Octavian M Machidon, Mihai Duguleana, and Marcello Carrozzino. 2018. Virtual humans in cultural heritage ICT applications: A review. *Journal of Cultural Heritage* 33 (2018), 249–260.
- [43] Hye Sung Moon, Grady Orr, and Myoungsoon Jeon. 2022. Hand tracking with vibrotactile feedback enhanced presence, engagement, usability, and performance in a virtual reality rhythm game. *International Journal of Human-Computer Interaction* 39, 14 (2022), 1–12.
- [44] Masahiro Mori, Karl F MacDorman, and Norri Kageki. 2012. The uncanny valley [from the field]. *IEEE Robotics & automation magazine* 19, 2 (2012), 98–100.
- [45] Niels Christian Nilsson, André Zenner, and Adalberto L Simeone. 2021. Propping up virtual reality with haptic proxies. *IEEE Computer Graphics and Applications* 41, 5 (2021), 104–112.
- [46] Gianpaolo Palma, Sara Perry, and Paolo Cignoni. 2021. Augmented virtuality using touch-sensitive 3D-printed objects. *Remote Sensing* 13, 11 (2021), 2186.
- [47] Isabel Pedersen, Nathan Gale, Pejman Mirza-Babaei, and Samantha Reid. 2017. More than meets the eye: The benefits of augmented reality and holographic displays for digital cultural heritage. *Journal on Computing and Cultural Heritage (JOCCH)* 10, 2 (2017), 1–15.
- [48] Sara Perry. 2019. The enchantment of the archaeological record. *European Journal of Archaeology* 22, 3 (2019), 354–371.
- [49] Eva Pietroni, Christie Ray, Claudio Rufa, Daniel Pletinckx, and Iefke Van Kampen. 2012. Natural interaction in VR environments for Cultural Heritage and its impact inside museums: The Etruscanning project. In *Proc. 18th International Conference on Virtual Systems and Multimedia*. 339–346.
- [50] David A Plecher, Christian Eichhorn, Khaled M Seyam, and Gudrun Klinker. 2020. ARsinoë-learning egyptian hieroglyphs with augmented reality and machine learning. In *2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 326–332.
- [51] Daniel Pletinckx. 2007. *Virtex: A multisensory approach for exhibiting valuable objects*. Interactive Institute AB.
- [52] Elizabeth Pye. 2016. *The power of touch: Handling objects in museum and heritage context*. Routledge.
- [53] Juno Rae and Lizzie Edwards. 2016. Virtual reality at the British Museum: What is the value of virtual reality environments for learning by children and young people, schools, and families. In *Proc. Conference on Museums and the Web*.
- [54] Priscilla Ramsamy, Adrian Haffegge, Ronan Jamieson, and Vassil Alexandrov. 2006. Using haptics to improve immersion in virtual environments. In *Proc. 6th International Conference on Computational Science*. 603–609.
- [55] Simone Restivo, Alberto Cannavò, Martina Terzoli, Davide Mezzino, Roberta Spallone, and Fabrizio Lamberti. 2023. Interacting with Ancient Egypt Remains in High-Fidelity Virtual Reality Experiences. In *Proc. EUROGRAPHICS Workshop on Graphics and Cultural Heritage*. 1–6.
- [56] José-Luis Rodríguez, Ramiro Velázquez, Carolina Del-Valle-Soto, Sebastián Gutiérrez, Jorge Varona, and Josué Enríquez-Zarate. 2019. Active and passive haptic perception of shape: Passive haptics can support navigation. *Electronics* 8, 3 (2019), 355.
- [57] Valentin Schwind, Katrin Wolf, and Niels Henze. 2018. Avoiding the uncanny valley in virtual character design. *interactions* 25, 5 (2018), 45–49.
- [58] Michael James Scott, Alcwyn Parker, Edward J Powley, Rob Saunders, Jenny R Lee, Phoebe Herring, Douglas Brown, and Tanya Krzywinska. 2018. Towards an interaction blueprint for mixed reality experiences in GLAM spaces: The augmented telegrapher at Porthcurno Museum. In *Proc. 32nd International BCS Human Computer Interaction Conference* 32. 1–5.
- [59] Paolo Sernani, Sergio Vagni, Nicola Falconelli, Dagmawi Neway Mekuria, Selene Tomassini, and Aldo F Dragoni. 2020. Voice interaction with artworks via indoor localization: A vocal museum. In *Proc. 7th International Conference on Augmented Reality, Virtual Reality, and Computer Graphics*. 66–78.
- [60] Orit Shaer, Eva Hornecker, et al. 2010. Tangible user interfaces: past, present, and future directions. *Foundations and Trends in Human-Computer Interaction* 3, 1–2 (2010), 4–137.

- [61] Adalberto L Simeone, Eduardo Velloso, and Hans Gellersen. 2015. Substitutional reality: Using the physical environment to design virtual reality experiences. In *Proc. 33rd Annual ACM Conference on Human Factors in Computing Systems*. 3307–3316.
- [62] Mel Slater, Bernhard Spanlang, and David Corominas. 2010. Simulating virtual environments within virtual environments as the basis for a psychophysics of presence. *ACM Transactions on Graphics* 29, 4 (2010), 1–9.
- [63] Charles Spence. 2016. Making sense of touch: A multisensory approach to the perception of objects. In *The Power of Touch*. Routledge, 45–61.
- [64] Jocelyn Spence, Dimitrios Paris Darzentas, Yitong Huang, Harriet R Cameron, Eleanor Beestin, and Steve Benford. 2020. VRtefacts: Performative substitutional reality with museum objects. In *Proc. ACM Designing Interactive Systems Conference*. 627–640.
- [65] Oliviero Stock. 2001. Language-based interfaces and their application for cultural tourism. *AI Magazine* 22, 1 (2001), 85–85.
- [66] William Swartout, David Traum, Ron Artstein, Dan Noren, Paul Debevec, Kerry Bronnenkant, Josh Williams, Anton Leuski, Shrikanth Narayanan, Diane Piepol, et al. 2010. Ada and Grace: Toward realistic and engaging virtual museum guides. In *Proc. 10th International Conference on Intelligent Virtual Agents*. 286–300.
- [67] Stella Sylaiou, Vlasios Kasapakis, Elena Dzardanova, and Damianos Gavalas. 2019. Assessment of virtual guides' credibility in virtual museum environments. In *Proc. 6th International Conference on Augmented Reality, Virtual Reality, and Computer Graphics*. 230–238.
- [68] Stella Sylaiou, Vlasios Kasapakis, Damianos Gavalas, and Elena Dzardanova. 2020. Avatars as storytellers: Affective narratives in virtual museums. *Personal and Ubiquitous Computing* 24 (2020), 829–841.
- [69] Diego Tamburini, Joanne Dyer, Marie Vandenbeusch, Matilde Borla, Debora Angelici, Maurizio Aceto, Cinzia Oliva, Federica Facchetti, Sara Aicardi, Patrizia Davit, et al. 2021. A multi-scalar investigation of the colouring materials used in textile wrappings of Egyptian votive animal mummies. *Heritage Science* 9, 1 (2021), 1–26.
- [70] Ana Martí Testón and Adolfo Muñoz. 2021. Digital avatars as humanized museum guides in the convergence of extended reality. In *Proc. Conference on Museums and the Web*.
- [71] Thomas J. Whelan. 2008. Social presence in multi-user virtual environments : A review and measurement framework for organizational research. <https://api.semanticscholar.org/CorpusID:15312067>
- [72] Bob G Witmer, Christian J Jerome, and Michael J Singer. 2005. The factor structure of the presence questionnaire. *Presence: Teleoperators & Virtual Environments* 14, 3 (2005), 298–312.
- [73] Bob G Witmer and Michael J Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence* 7, 3 (1998), 225–240.
- [74] Ryan Yung, Catheryn Khoo-Lattimore, and Leigh Ellen Potter. 2021. Virtual reality and tourism marketing: Conceptualizing a framework on presence, emotion, and intention. *Current Issues in Tourism* 24, 11 (2021), 1505–1525.
- [75] Can Zhao, Michael B Twidale, and David M Nichols. 2018. Acquiring metadata to support biographies of museum artefacts. In *Proc. International Conference on Asian Digital Libraries*. 304–315.