



Usability, user experience and mental workload in a mobile Augmented Reality application for digital storytelling in cultural heritage

Lucio Tommaso De Paolis¹ · Carola Gatto² · Laura Corchia² · Valerio De Luca¹

Received: 13 November 2021 / Accepted: 17 October 2022
© The Author(s) 2022

Abstract

Augmented Reality (AR) has become an increasingly used technology to support and enhance the enjoyment of cultural heritage. Particularly relevant is its importance for digital storytelling: by framing a portion of a fresco or painting with a smartphone, an AR mobile application can provide contextually relevant information, also in the form of multimedia content, that can help the user to understand the story and meaning behind the images. In this type of application, human factors are of fundamental importance for the effectiveness of the narrative: a mobile AR application must avoid distracting the user's attention from the content in order to encourage a good level of concentration and immersion. The case study presented in this paper deals with a mobile AR application developed to guide visitors in the interpretation of the frescoes inside the Basilica of Saint Catherina of Alexandria in Galatina. The aim of the study is the analysis of the relations among usability, user experience and mental workload factors in AR-based digital storytelling.

Keywords Augmented Reality · Mobile · Digital storytelling · Cultural heritage · Usability · User experience · Mental workload

1 Introduction

In recent years, the evolution of technologies has opened up new frontiers and found innovative solutions in the field of tourism and enhancement of cultural heritage (Bekele et al. 2018). Augmented Reality (AR) is often adopted to complement traditional forms of fruition and offer visitors new ways of exploring and acquiring knowledge during the visit of a place of historical and artistic interest. Based on

the approach of edutainment (Paolis et al. 2011a, b), which exploits game-oriented environments (Paolis et al. 2010) and multichannel and multisensory platforms (Paolis 2013) for educational purposes, AR fosters a direct involvement of the visitors (Cervellini and Rossi 2011), by providing them with relevant and contextual information. In museums and archaeological sites, it allows to advance reconstructive hypotheses about the structural parts, polychromy and sculptural furnishings no longer present on site. The visitor, using a tablet or smartphone, can thus take a “trip back in time” and see the differences between what was and what is no longer there. This is particularly useful for environments that have changed considerably over time or mutilated sculptures and paintings whose polychromy has been lost. AR applications also represent an alternative tool to the traditional guided tour, enabling a deeper knowledge and understanding of the artwork, through an interactive guide based on a stimulating, immersive and engaging communicative impact.

In this context, what might partially disorient the user stems from the fact that AR establishes a connection between a story that follows a narrative order and real environments that can be explored in a nonlinear manner

✉ Valerio De Luca
valerio.deluca@unisalento.it

Lucio Tommaso De Paolis
lucio.depaolis@unisalento.it

Carola Gatto
carola.gatto@unisalento.it

Laura Corchia
laura.corchia1@unisalento.it

¹ Department of Engineering for Innovation, University of Salento, Lecce, Italy

² Department of Cultural Heritage, University of Salento, Lecce, Italy

(Shilkrot et al. 2014). Also in the light of this consideration, this work starts from a mobile AR application developed for the promotion of Basilica of Saint Catherina of Alexandria in Galatina, introduced in a previous work (Cisternino et al. 2021), to conduct a deeper study aimed at investigating the interdependencies among usability, user experience factors and mental workload from users' answers to SUS, UEQ and NASA-TLX standard questionnaires. At the end of the analysis, a possible integration of these human factors into a unified framework is also discussed.

In particular, some important research questions addressed by this work are:

- Is the attractiveness of an AR application such as the one analysed in this paper influenced in any way by usability?
- Is there a significant correlation between the novelty effect and the attractiveness of the application?
- Is it mentally more demanding to use the application or to learn how to use it?
- Does the storytelling influence the mental workload?

The rest of the paper is structured in this way: Sect. 2 presents the related work about AR applications for cultural heritage and some models for usability and user experience evaluation; Sect. 3 introduces digital storytelling; Sect. 4 briefly describes the case study on which the AR application presented in Sect. 5 focuses; Sect. 6 introduces the experimental methodology and the questionnaires used for the post hoc analysis; Sect. 7 analyses the collected data about users' impressions; Sect. 8 discusses and summarizes the main findings; and Sect. 9 concludes the paper.

2 Related work

Several AR applications have been developed especially in the archaeological field, being this technology able to offer the user the possibility to better understand the transformations occurred over time through the visualization of faithful reconstructions of monuments and landscape (Bonacini 2014). The European Union has immediately grasped the potential offered by this technology and has funded several cultural projects, such as ARCHEOGuide (Augmented Reality based Cultural Heritage On-Site Guide) (Vlahakis et al. 2002) and iTACITUS (Intelligent Tourism and Cultural Information through Ubiquitous Services) (Zoellner et al. 2007). The former is an application that guides the user during the visit of an archaeological site using a laptop equipped with an HMD visor that returns the image of the temple of Zeus at Olympia, superimposing it on the existing ruins. The latter is a tool designed to enhance the points of interest of a territory by providing information in AR simply by pointing a device towards an artwork.

Another mobile AR application (Cisternino et al. 2018) was developed to support the promotion of the archaeological areas of the "Castello di Alceste" Diffuse Museum in San Vito dei Normanni (Brindisi) and the "Fondo Giuliano" site in Vaste (Lecce): when an aerial photograph of the areas is framed using a smartphone, 3D models and other contextual information are visualized.

The ArkaeVision project (Bozzelli et al. 2019) proposed an integrated VR/AR framework for a gamified exploration with elements of digital fiction and an engaging storytelling: it employs VR technology for the exploration of the Temple of Paestum and AR technology for the exploration of the slab of the Swimmer Tomb.

Another work (Paolis et al. 2021) studied the use of virtual and mixed reality technologies for the promotion of an underground oil mill, addressing place accessibility issues and the valorization of itineraries and rural heritage.

An outdoor location-based mobile AR application that allows users to interact with both virtual and physical objects was presented in Ping et al. (2020): experimental tests were conducted to assess navigation accuracy, ease of use, interaction naturalness, attraction, immersion and users' attitude towards the application. The tests revealed also the ability of the application to improve the learning effectiveness and extend the focusing time of participants.

A special kind of on-site exploration made possible by mobile AR is based on virtual portals, located in specific points of interest, that emulate a transition to a virtually reconstructed past reality when users pass through them (Cisternino et al. 2019).

Location-based AR can also exploit unmanned aerial vehicles (UAVs) to take pictures of an archaeological site from an aerial perspective and display various contextual information (3D models, textual information, etc.) over them (Botrugno et al. 2017).

AR has also been used in various museum. The Museum of London was one of the first museum that makes use of AR applications: by launching StreetMuseum, an application that uses geolocation to render certain views of the city in AR, the user can take a trip back in time and sees monuments as they appeared in the past. Another museum that was able to take advantage of AR technology to enhance the understanding of its collections was the Franklin Institute in Philadelphia, which, from 30 September 2017 to 4 March 2018, hosted the Terracotta Warriors of The First Emperor exhibition (Terracotta 2021), including ten of the 8,000 famous statues unearthed in 1974 by a Chinese farmer. The AR application that accompanied the exhibition allowed to frame the warriors and to digitally reconstruct their weapons. By zooming in, it was also possible to admire the statues in their details. In Italy, too, several museum used this technology to enhance cultural heritage, transforming the traditional guided tour into a more interactive and engaging

experience. Among the regions that first proposed cultural tourism routes based on mobile applications, we would like to point out Tuscany and Apulia, whose applications are configured as interactive guides of the territory, with some maps indicating the points of interest, the routes to reach them and some additional information that enriches the knowledge. The potential AR interaction patterns for guided tours in museums were explored in Liu et al. (2021), where a possible combination of handheld device and head-mounted display was studied.

The mobile AR application described in Paolis et al. (2018) recognizes the sketches on the Atlantic Codex and superimposes animated 3D models showing the structure and the working principles of the machines designed by Leonardo Da Vinci. Another application based on touchless interaction using the Kinect device was made to allow further study of these machines and tests were carried out to assess the impact in a learning context (Paolis et al. 2019).

Other applications have instead focused on a single artwork, allowing the user to travel back in time through 3D reconstructions usable by means of visors or permitting to deepen the executive technique, penetrating into the brushstrokes or preparatory layers of a painting and, in some cases, returning the original colours now lost. An example of this is the innovative Project “L’Ara com’era”, realized with the scientific collaboration of the Superintendence of Roma Capitale and aimed at enhancing the Ara Pacis of Augustus. Visitors, wearing a Samsung Gear VR visor inside which is placed a smartphone, have the opportunity to review the colours that originally enriched the monument and to view 3D movies that tell some aspects. Even the Hall of Frescoes in the Palazzo Comunale of Tarquinia (Hall 2021) has been the subject of an interesting AR project that has given voice to the characters depicted, allowing them to tell the story, anecdotes and curiosities about the past of the city. There are also museums created specifically to take advantage of AR technology, as in the case of MAUA—Museo di Arte Urbana Aumentata (Museum of Augmented Urban Art) (MAUA 2021), a widespread museum created in Milan, Palermo and Turin. The street artworks, once framed with the smartphone, come to life and transform.

A different kind of AR applications applied to cultural heritage is the Spatial Augmented Reality (SAR), also mentioned as video mapping (VM), a particular form of AR that consists in the projection of light beams on surfaces in order to transform the facades of buildings into screens, altering the real vision and enriching it with content (Bimber and Raskar 2005; Cisternino et al. 2021; Paolis et al. 2022).

A multiplatform usability evaluation test, based on interviews, observations, think-aloud protocol and questionnaires, was conducted on a digital storytelling application for the city of Timisoara in Romania employing interactive touchscreen table, desktop/laptop, mobile and AR platforms

(Vert et al. 2021). The study demonstrated the strengths of various types of technologies in cultural heritage: an interactive touchscreen table fosters social interaction, a desktop/laptop application allows for detailed exploration of content, a mobile application allows for on-the-fly exploration and sharing of information with other users and an AR application offers the possibility of deepening topics at certain landmarks.

The study in Jin et al. (2022) compared natural user interface and graphical user interface for a Hololens narrative application: between the two modes, the former turned out to be the best performing system for users without Role Playing Game (RPG) experience, while the latter turned out to be the best performing system for users with RPG experience.

2.1 Usability and user experience

According to ISO 9241-11 International Organisation for Standardisation (2018), usability is “the extent to which a system, product or service can be used by specified users to achieve specific goals with effectiveness, efficiency and satisfaction in a specified context of use”.

User performance is influenced by mental workload (Cain 2004), also known as cognitive workload, which is the mental effort needed to perform the tasks. It can be evaluated by means of quantitative performance tests, physiological measures or subjective feedback collected through questionnaires (Moustafa et al. 2017).

Various questionnaires were developed to be administered to users after their experience with a system or application to assess the usability they perceived (Assila et al. 2016; Lewis 2018; Hajesmaeel-Gohari and Bahaadinbeigy 2021).

The HARUS questionnaire was developed for the usability of mobile AR applications (Santos et al. 2014, 2015): it was conceived from an analysis of perceptual and ergonomic issues, which inspired respectively the dimensions of comprehensibility, i.e. the ease of understanding the content, and manipulability, i.e. the ease of handling the device during a task.

User experience goes beyond the traditional concept of usability and covers additional factors such as usefulness, emotional factors and design elegance (Vosinakis and Koutsabasis 2018). As stated by ISO 9241-210 International Organisation for Standardisation (2019), it deals with “perceptions and responses that result from the use or anticipated use of a product, system or service”.

The Augmented Reality Immersion (ARI) questionnaire (Georgiou and Kyza 2017) was designed to measure immersion in location-based AR. It is made up of 30 items, which can be grouped into six factors: *interest*, *usability*, *emotional attachment*, *focus of attention*, *presence* and *flow*.

The framework proposed in Okanovic et al. (2022) for extended reality environments consists of an introductory

part, a quantitative evaluation and open questions. It covers three subscales that can be mapped with two constructs of the UTAUT model introduced several years earlier (Venkatesh et al. 2003): *immersion* and *edutainment* are linked to *performance expectancy* (i.e. how much an individual believes that using the system will help him/her improve his/her performance), while perceived *ease of use* coincides with the definition of *effort expectancy*. Data analysis revealed a positive effect of the narrative quality on *immersion* and *edutainment*, even when ease of use issues occur.

Among the various models developed for storytelling evaluation, the Narrative Engagement Scale (Busselle and Bilandzic 2009) consisted of several items grouped into four dimensions:

- Narrative understanding;
- Attentional focus;
- Emotional engagement, which refers both to the ability to feel the emotions of the characters and to feelings towards them;
- Narrative presence, which refers to a transition from the real world to the story world.

The emotional gratification scale presented in Bartsch (2012) consists of three factors describing rewarding feelings (fun, thrill and emphatic sadness) and four factors concerning emotions in the context of social and cognitive needs (contemplative emotional experiences, emotional engagement with characters, social sharing of emotions and vicarious release of emotions).

Jin et al. (2022) designed a scale to assess the empathy and the heartfelt connection to the protagonist by combining an emotional engagement scale extracted from (Bartsch 2012) and a contemplativeness scale derived from (Bartsch et al. 2014; Bartsch 2012).

A comprehensive framework for the evaluation of the impact of interactive guides and AR/VR technologies on museum settings is the MUSETECH model (Damala et al. 2019): it considers three symbolic entities (the Visitor, the Cultural Heritage Professional and the Museum) and divides the life cycle of technologies into four phases (design, content, operation and compliance).

The questionnaire used in Boskovic et al. (2017) for a cultural heritage application consists of a user profiling section, 30 items for the evaluation of immersion, edutainment and usability and some final open questions to collect users' opinions on their favourite and most problematic parts. The results of the application to a case study revealed the potential of virtual models to engage users in the exploration of historical artefacts by increasing their motivation, even in the presence of navigation and visibility issues. They highlighted also the importance of the

quality of the production and the performance of digital stories for a high level of user immersion.

3 Digital storytelling

According to Abbott (Porter Abbott 2014), “story is an event or sequence of events”, while “narrative discourse is those events as represented”. Ryan (2006) also included other elements in the narrative concept such as a story world, intelligent characters, a timeline and meaningful events.

Digital storytelling refers to a variety of digital media platforms designed for narrative purposes (Miller 2019). In interactive digital storytelling, the user plays an active role, by influencing the flow or even the content of the story (Okanovic et al. 2022).

According to the model proposed in Roth and Koenitz (2016), user experience in interactive digital narratives can be described through 12 dimensions clustered into three experimental qualities: *agency*, *immersion* and *transformation*.

Agency, defined as the ability to take meaningful actions (Murray 1997), includes *usability*, which is a precondition for any enjoyable user experience and significantly influences *effectance*, *autonomy* and *user satisfaction*. In this context, *effectance* is the effect of an action, while *autonomy* refers to the freedom to choose from various options that can influence a narrative without feeling constrained in one direction. However, it is important to underline that high availability of options does not necessarily produce a more enjoyable experience.

Murray (1997) describes narrative as a “transformational” experience that can change the user. Interactive Digital Narratives foster a more direct connection (Murray 1997) through an active participation of the user, also by providing alternative paths and outcomes in different sessions.

An important factor of transformation is *eudaimonic appreciation*, which represents a special kind of engagement that derives from the aesthetic presentation of the interactive narration and leads the user to develop a personal dimension linked to his/her previous experiences. The distinction between *hedonistic enjoyment* and *eudaimonic appreciation* was explored in Oliver and Raney (2011): the former is related to pleasure and amusement, while the latter is related to the need to research and reflect on the meaning, truths and purposes of life.

Other components of the transformation dimension are *affect*, which consists of various measures of the affective states perceived by the user, and *enjoyment*, which represents entertainment in a broader sense (Roth and Koenitz 2016).

3.1 Augmented Reality for digital storytelling

One form of interactive digital storytelling is provided by Augmented Reality (AR), a technology that attempts to blur the line between reality and fiction (Shilkrot et al. 2014). A widely discussed topic is the connection that AR establishes between the story world, which follows a narrative model, and the real world, which the user can freely explore in a nonlinear way (Shilkrot et al. 2014). Moreover, AR exploits the concept of minimal departure, where a fictional story world can be understood due to its derivation from a real world that users are already able to perceive for themselves (Herman et al. 2007).

Based on interaction mode, AR-based narratives can be classified into three categories (Shilkrot et al. 2014):

- Point of view-based exploration, which involves the user in a first-person game experience, where they are able to delve into the story through the eyes of a character;
- Space-based exploration, where the user can activate the playback of storytelling content associated with frame landmarks or markers in a real-world space;
- Ontological interaction, which has the possibility to alter the plot or the world of AR narrative.

According to the type of experience in which users are involved, AR-based narratives can be divided into three categories (Shilkrot et al. 2014):

- Situated augmented narratives, which typically have a local nature and take place within a limited time;
- Location-based narratives, which can augment a wider area of the physical world by exploiting various portals that give access to a story world;
- World-level narratives, which extend globally and over a long period of time.

This paper examines the case of a space-based exploration with a location-based narrative, made possible by an AR application that is articulated through eleven points of interest associated with portions of the frescoes located at various points in the Basilica (as explained in detail in Sect. 5).

4 The case study: the Basilica of Saint Catherine of Alexandria in Galatina

After the birth of the Kingdom of Italy, the Basilica of Saint Catherine of Alexandria in Galatina (Lecce, Italy) was classified as a “National Monument of the first category” and in 1929 it was entrusted to the Friars Minor of the Province of St. Joseph. At the end of the last century, it obtained the

nomination of Pontifical Minor Basilica. The facade, interested by a recent restoration that brought back to light the warm colour of the stone, is characterized by the presence of some disharmonious elements caused by the different interventions that followed one another during the centuries. A stringcourse frame runs horizontally, making the lower part more projecting than the upper register, which houses, in the centre, a stone rose window. The three portals correspond to the main nave and to the two lateral ambulatories.

The central portal, larger than the other two, is characterized by the presence of a porch resting on two columns supported by lions and surmounted by griffins acephalous. The three bands of stone that frame the wooden portal are richly carved with acanthus spirals and anthropomorphic figures, zoomorphic and phytomorphic. On the architrave is placed the relief representing Christ among the twelve apostles, according to a frontal and hierarchical disposition that denounces the presence of archaisms.

The rose window placed in the upper curtain wall is decorated with two orders of concentric circles carved with an ornate plant motif. The twelve rays welcome the stained glass windows and converge towards a central polychrome oculus representing the coats of arms of the d’Anjou Durazzo and d’Enghien-Brienne.

The interior of the basilica is striking for its solemn longitudinal layout and for the pictorial decoration that covers its walls (Fig. 1). The central nave is divided into three spans with cross vaults and is divided from the lateral spans by three lowered pointed arches. The spans are marked by transverse ogival arches resting on capitals sculpted with ornamental motifs taken from the Romanesque repertoire.

The pictorial decoration, originally, had to cover all the walls of the building. Today, the walls of the church preserve traces of two pictorial campaigns: the first one, which took place around 1391, is visible in the lower registers where devotional frescoes depicting saints appear. The second one, of uncertain dating (Cuciniello 2014), unfolds on the walls and on the vaults. In the first span and in the counter-façade are depicted scenes from the Apocalypse of John. On the vault are painted the Virtues. The second span is dedicated to the stories of Genesis, overlapped by the representation of the Sacraments. The walls of the third bay are occupied by Christological scenes from the Gospel, while the vault is decorated with the angelic hierarchies.

The story of the life of Saint Catherine is narrated in the presbytery which, at the top, is dominated by representations of the Evangelists and the Doctors of the Church. In the right aisle, instead, where the Orsini chapel is located, there are frescoed scenes of the life of Christ and the Virgin. Traces of a previous pictorial campaign, probably dating back to 1391, are recognizable in the lower parts of the basilica where there are numerous votive representations. According to the most recent criticism, different artists dedicated themselves

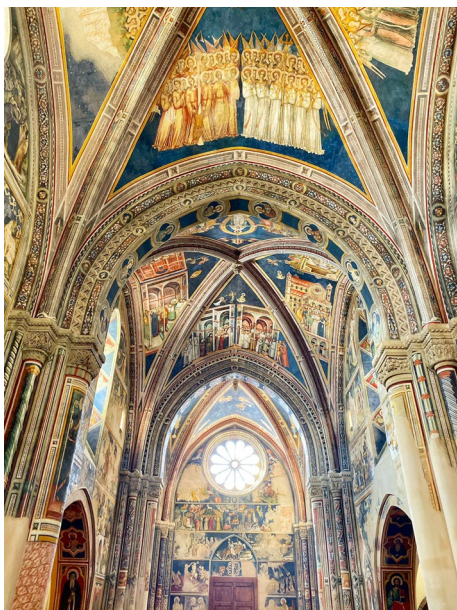


Fig. 1 Interior of the Basilica of Saint Catherine of Alexandria in Galatina (courtesy of the Archdiocese of Otranto)

to the Galatian pictorial cycle because of their formation and origin (Casciaro 2017). Scholars have recognized the intervention of three workshops headed by as many masters: one worked on the stories of Saint Catherine and on the vaults with the Evangelists, the angelic hosts and the Sacraments; another seems to have taken care of the walls of the second and third span, with stories from Genesis and the life of Christ; a third group of artists is active in the cycle of the Apocalypse. If the first two artists reveal a Venetian education, the master of the Apocalypse seems to have had knowledge of Neapolitan painting, as attested by the similarity with Giotto's lost cycle in Santa Chiara. The hand of a fourth master, more inclined to anecdotal narration, can instead be identified in the frescoes that cover the Orsini Chapel.

5 The mobile AR application

5.1 Design

The AR application was born from the need to provide visitors with an innovative tool that can complement the traditional guided tour and enhance its use through thematic tours capable of enriching the cultural offer and facilitate the reading of the valuable pictorial cycles that cover the walls of the building. The AR application can offer the user a more engaging and educational experience, providing him with a wealth of information conveyed through multimedia content.

An in-depth preliminary study on the historical and artistic aspects of the Basilica was carried out to identify the pictorial portions of greatest interest: in particular, an accurate bibliographical study and numerous inspections were necessary to identify the position of the frescoes and the sources of illumination. The main challenges were the construction of a storytelling based on solid historical critical foundations and the selection of all those pictorial portions that could be easily framed by the users' devices. The iconographic and iconological heterogeneity of the cycles, whose interpretation is still the subject of studies by art historians, has led to the selection of eleven points of interest, reported in Tables 1 and 2, chosen on the basis of historical and stylistic peculiarities little known to most visitors. The constructed itinerary is guided by the voice (interpreted by a speaker) of Pietro Cavoti (1819 - 1890), an illustrious artist from Galatina who for a long time waited for the studies of the Basilica in his capacity as President of the Conservative Commission of the Monuments of Terra d'Otranto and as Inspector of the Monuments (Montinari 1978). He, in fact, had made numerous drawings and sketches, accompanying them with historical and stylistic annotations. In order to better frame his activity, in-depth research was carried out at the Fondo Cavoti that allowed the photographic acquisition of numerous watercolours and pencil sketches.







Even though several tests revealed the possibility to recognize also the pictorial portions placed in the upper registers and on the vaults, it could be uncomfortable for the visitor to enjoy the content while keeping his arms stretched upwards. For this reason, in the design of the application, it was decided to keep the AR content even when the visitor moves his device along other directions. Moreover, since visitors are not allowed to access the presbytery areas, where the Funeral Monuments and the cycle with the Stories of St. Catherine are located, such frescoes are excluded from the list of points of interest.

5.2 Interface

A careful design of the graphic interface was carried out to make the application pleasant and easy to use even for inexperienced users, in order to meet the principles of effectiveness, efficiency and user satisfaction. A good interface and an effective storytelling will provide a good basis for experimentation for the analysis of the relationships between various user experience indicators that will be conducted in the following sections.

In particular, the application interface is made up of graphic shapes taken from architectural and pictorial elements and colours with a strong symbolic value, such as white and gold, which are a clear reference to the royalty of the Basilica's patronage.

Table 1 Points of interest of the guided tour—part 1

Point of interest	Pictorial cycle	Fresco's title	Audio guide content
			Pietro Cavoti (1819–1890) guides the visitor along the entire route. At this point, he introduces himself, tells about his activity inside the Basilica and invites the user to follow him along the itinerary
	Cycle of Revelation	Noble coats of arms	This fresco depicts the noble coats of arms of the patrons. Pietro Cavoti tells the story of the construction of the Basilica and the journey made by Raimondo Orsini del Balzo to Mount Sinai where Saint Catherine of Alexandria is buried. According to legend, he bit off the finger of the Saint and brought it as a relic to his Christian temple
	Cycle of revelation	The woman riding the dragon	This episode is taken from the Revelation of John. Pietro Cavoti introduces the pictorial cycle to which it belongs and explains to the user how the episodes were depicted by the artist
	Cycle with stories from Genesis	The construction of Babylon	In this panel, the anonymous artist has depicted the construction of Babylon by setting the scene in a mediaeval shipyard. After pointing out the mediaeval tools and machines in detail, Pietro Cavoti tells how churches were built in the Middle Ages
	Cycle with stories from Genesis	The universal deluge	In the fresco depicting the Universal Flood, the painter depicted the Ark as a wooden box. Pietro Cavoti points to the realism of the bodies of the dead floating in the water, swollen and bruised
	Triumph of the Church and the Seven Sacraments	The Sacrament of Matrimony	The scene of the Sacrament of Marriage is an opportunity for Cavoti to return to tell the story of Maria d'Enghien, who commissioned the cycle of frescoes together with her husband Raymond. According to some hypotheses, the bride with the crown depicted by the artist could be a portrait of the same client, who assumed the title of queen after having married the king of Naples Ladislaus of Durazzo in her second marriage (Vacca 1955)

After downloading the application, the user accesses the interface in Fig. 2, which includes a full-screen photograph of the main nave of the Basilica, the application logo and a button that allows access to instructions. At the bottom, the arrow on the right gives access to a plan of the basilica containing clickable points of interest. A drop-down menu at the top provides access to textual information on the history, the pictorial cycles and the adjacent museum.

The map at the bottom left of Fig. 2 indicates the hot spots where the frescoes to be framed with the smartphone are located. The user starts the journey from point 1 and gradually moves towards the other points accompanied by the audio guides. By clicking on a point, the user accesses a preview screen that helps him understand which portion of the painting to frame with the camera to activate the visualization of the AR content associated with it. It contains a photographic reproduction, the title of the fresco and indications regarding the cycle to which it belongs. The camera-shaped button at the bottom right gives access to the AR scenes: it is sufficient to frame the painting of interest in order to see the icons corresponding to the audio guides and, in some cases, to the virtual restoration of lost parts of frescoes or to the game that allows you to listen to the sound of some mediaeval musical instruments depicted in the frescoes.

5.3 Implementation

The AR application, developed in Unity (Unity 2021) for Android and iOS platforms, is based on the Vuforia development kit (PTC 2021) (version 2018 4.12f1): it supports a variety of 2D and 3D target types, enables markerless and multitarget feature tracking, and has an internal fiducial marker system known as VuMark. Additional features of the SDK include Occlusion Detection, Extended Tracking, and Cloud Storage, i.e. online space for storing images and evaluating them in terms of tracking quality.

The analysis of the context has led to choose a markerless tracking based on natural features and on the recognition of two-dimensional images represented in the frescoes of the Basilica. Therefore, a photographic campaign was carried out to obtain the application targets. At the end of each inspection, the whole material was acquired and displayed, and in order to improve the reading of the lenses by Vuforia, each shot was subjected to a post-production through PhotoShop aimed at cropping the images, increasing the contrasting colour and keeping them within the maximum limit of 2 MB imposed by Vuforia. A subsequent evaluation of the photographs allowed us to assess how suitable each image is for tracking: the higher the number of discontinuities, i.e. features, the more suitable that photograph will be for image recognition.

6 Methodology of analysis and experimentation

The factors of usability and user experience and the connections between them have already been discussed in the literature: most of the published papers tried to define theoretical models or questionnaires used to individually assess the characteristics of an application. At the same time, several models have been proposed to describe various factors that characterize digital storytelling and its effectiveness. This paper follows a different approach, which aims to evaluate by means of a post hoc analysis the connection between different usability and user experience factors in the specific case of a mobile AR application for cultural heritage, also assessing possible correlations with the mental workload.

The tests involved 41 visitors of the Basilica of various ages (from 23 to 77, distributed as illustrated by the density plot in Fig. 3) without previous experience with AR technologies.

As the Basilica is a sacred place, they were provided with both tablets and headphones at the beginning of the visit to allow them to enjoy the contents of the audio guides in silent mode.

After the visit, visitors were asked to fill in a questionnaire made up of items taken from NASA-TLX (Nasa and Administration 2010), SUS (Brooke 1996) and UEQ (Laugwitz et al. 2008) standard tools dealing with the perceived workload, the application's usability and the user experience.






6.1 NASA-TLX

The NASA Task Load Index (NASA-TLX) (Nasa and Administration 2010) was designed to assess the perceived workload during the execution of a task. Among the 6 NASA-TLX items, physical demand and temporal demand were discarded, since the only physical activity required to use the application consists in touch-based interactions on a tablet and no hurry is imposed during the tasks. The considered items were *Mental demand* (how much thinking, deciding or calculating was required to perform the task), *Performance* (the success in accomplishing the application tasks), *Effort* (how hard the user had to work to achieve his/her level of performance) and *Frustration*, with scores expressed on a scale from 0 to 6.

6.2 SUS

The System Usability Scale (SUS) (Brooke 1996; Borsci et al. 2009) consists of a 10 item questionnaire with five response options ranging from "Strongly Disagree" to "Strongly Agree".

Table 2 Points of interest of the guided tour—part 2

Point of interest	Pictorial cycle	Fresco's title	Audio guide content
	Cycle with stories from the life of Christ	The Baptism of Christ	Pietro Cavoti points out a detail to the visitor: the young man in green clothes who appears in the scene would be, according to recent studies, Giovanni Antonio del Balzo Orsini.
	Cycle with Stories from the life of Christ	The Resurrection of Lazarus	Here again, the focus is on the realism of the painter who depicted the bystanders with their noses plugged to protect themselves from the bad odours emanating from the putrefaction of Lazarus' body.
	Cycle with Stories from the life of Christ	The ascent to Calvary	This scene, of which only the sinopia remains, is an opportunity to describe how a fresco was made in mediaeval times.
	Orsini Chapel pictorial cycle	The Coronation of the Virgin	Two angels appear in the scene playing some instruments. Pietro Cavoti points out that in the Basilica are represented 42 mediaeval musical instruments: waffles, double flutes, drums, strings, vielle, lutes, citolas, harps form the oldest illustrated encyclopaedia of instruments in Europe (Castaldo 2006). This looks to be a real iconographic treasure.
	Orsini Chapel pictorial cycle	Birth of the Virgin	Pietro Cavoti focuses attention on the domestic environments painted by the artist.

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

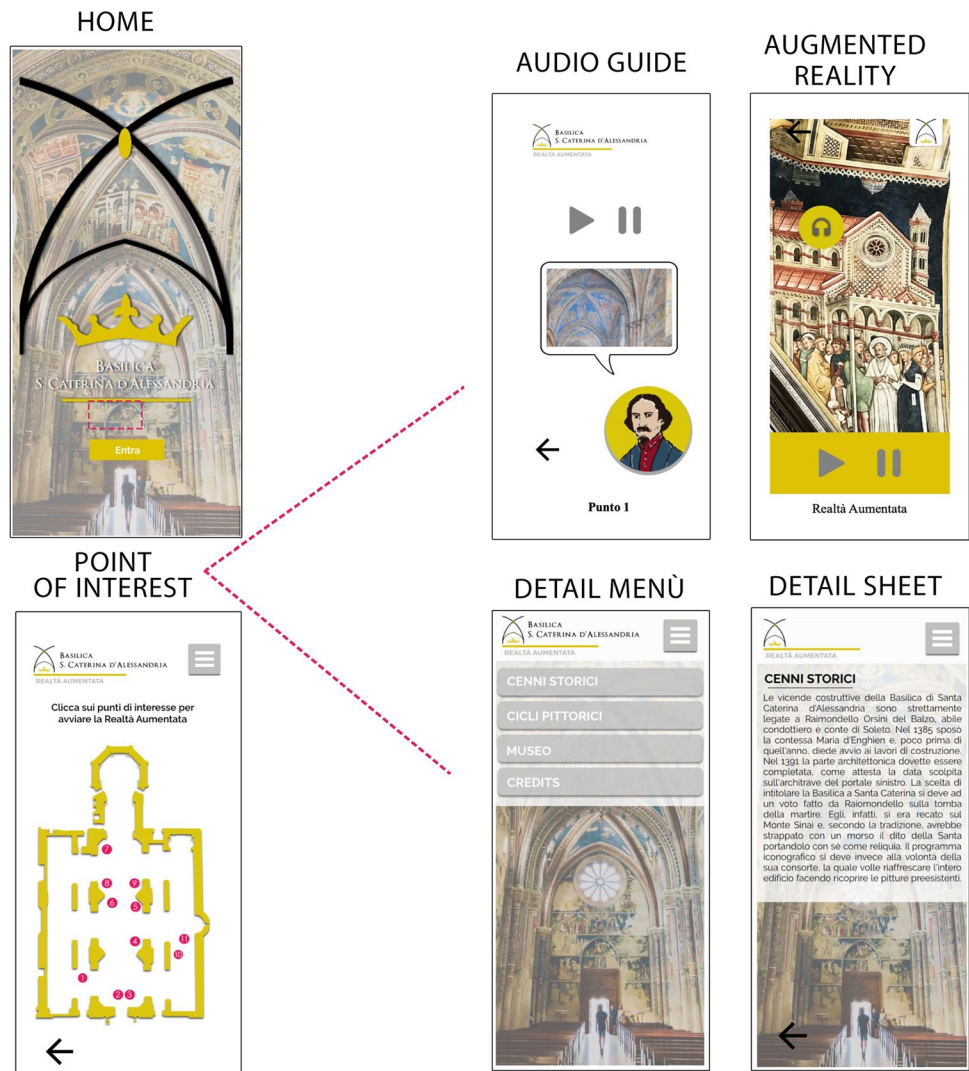
The SUS questionnaire is known to have a two-factor structure (Lewis and Sauro 2009), consisting of *Learnability* and *Usability* in a strict sense, which are weakly correlated (Borsci et al. 2009). The former is made up of items 4 and 10, while the latter is made up of all the remaining items.

6.3 UEQ

The 26 items of the UEQ questionnaire are grouped into 6 components (Laugwitz et al. 2008):

- *Attractiveness*, which is the general impression towards the product;
- *Efficiency*, which describes how fast and efficient the application is to use, including the organization of the user interface;
- *Perspicuity*, which is the ease to get familiar with the product;
- *Dependability*, which is the perception of controlling the interaction;
- *Stimulation*, which concerns the interest and motivation fostering the use of the application;
- *Novelty*, which expresses how creative and eye-catching the design is considered.

Fig. 2 AR application interface



7 Result analysis

Principal component analysis was the first method employed to study data dimensionality and detect any relation among the considered variables. Then a cluster analysis was carried out by means of the ICLUST package, based on a hierarchical algorithm (Revelle 1978). ICLUST and principal component analysis can provide different results, as they are based on two different approaches (Cooksey and Soutar 2006): while the former tries to maximize internal consistency and homogeneity, the latter aims at maximizing variance by considering all the items simultaneously. Internal consistency represents how closely related a set of items are. A well-known internal consistency measure is Cronbach's *alpha* (Cronbach

1951), which is defined as the mean of all the possible split-half reliabilities of a scale. However, this measure assumes a single underlying general factor. For this reason, Revelle introduced coefficient *beta* (Revelle 1979), defined as the minimum value among all the possible split-half reliabilities, to assess the scale homogeneity, which is assumed by coefficient *alpha*.

7.1 Principal component analysis

Principal component analysis (PCA) aims at grouping together multiple variables according to their variability and reducing the dimensionality of the data set. In the following subsections, principal component analysis was first applied to the items of each individual questionnaire, then to the factors describing the user experience in the UEQ

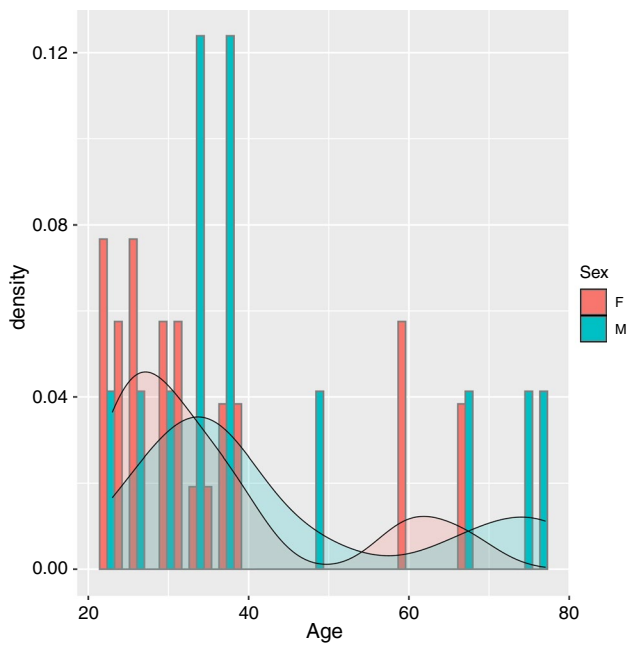


Fig. 3 Density plot of the ages of the users involved in the test

questionnaire and finally to the factors of all three questionnaires taken together.

7.1.1 PCA of NASA-TLX items

The plot in Fig. 4 represents the relationships among NASA-TLX items in a space made up of the first two dimensions (which account for the 75% of the total variability), since variables correlated with the first two principal components are the most important to explain the variability in the data set: the longer an arrow, the better the variable it refers to is represented in the factor map; moreover, positively related variables, such as *Effort* and *Frustration*, are represented by very close arrows, while negatively related variables, such as *Mental demand* and *Performance*, are represented by arrows pointing in opposite directions. In the same chart, two concentration ellipses enclose the clusters detected in the data set.

The bar plot in Fig. 5 depicts the square cosine (\cos^2) of NASA-TLX variables on the first two dimensions: *Performance* and *Mental demand*, which are the variables with the highest values (close to 1), are best represented by the first two principal components; on the other hand, lower \cos^2 values suggest a weaker representation of *Effort* and *Frustration* variables on the two components. Variables

that are correlated with the first two principal components are the most influential for the variability in the data set.

Table 3 shows that the first two and the first three components account for the 75% and the 93% of the total variability.

In principal component analysis, various types of rotation can be applied to the loading matrix to achieve the so-called simple structure, which should make their interpretation easier. After an initial formulation by Thurstone (1947), based on five criteria, a simpler formulation is now accepted, proposed by Kline (2002), according to which “each component should have a few high loadings with the rest of the loadings being zero or close to zero”. In this study *varimax* orthogonal rotation and *promax* rotation were considered: the former aims at maximizing variance among the squared values of loadings of each component (Kaiser 1958), while the latter is an oblique enhancement of *varimax* aimed at meeting the “simple structure” to a greater degree (Hendrickson and White 1964). After an orthogonal rotation component variances get changed, but components remain uncorrelated and variable communalities (which represent the part of the variance shared with other variables) are preserved.

The loadings in Table 4 were obtained by performing *varimax* and *promax* rotations on the first two components. The sign of a loading indicates whether a variable and a principal component are positively or negatively correlated. Only loadings higher than 0.4, highlighted in bold, were considered as significant. They revealed that the first component can be expressed as a combination of *Mental demand* and *Performance*: the opposite sign of the loadings indicates that *Performance* decreases as *Mental demand* increases. The second component can be expressed as a combination of *Effort* and *Frustration*, which are both negatively correlated with it.

The loadings in Table 5 were obtained by performing *varimax* and *promax* rotations on the first three components. The first component is still a combination of *Mental demand* and *Performance*, but now the second principal component coincides almost exclusively with *Effort*, while the third coincides with *Frustration*.

7.1.2 PCA of SUS items

The plot in Fig. 6 represents the relationships among SUS items in a space made up of the first two dimensions: they account for the 56.8% of the total variability and thus provide a weak representation, as confirmed by the short length of most of the arrows. SUS items 3 and 7 are represented by almost coinciding arrows: this would suggest a close relation

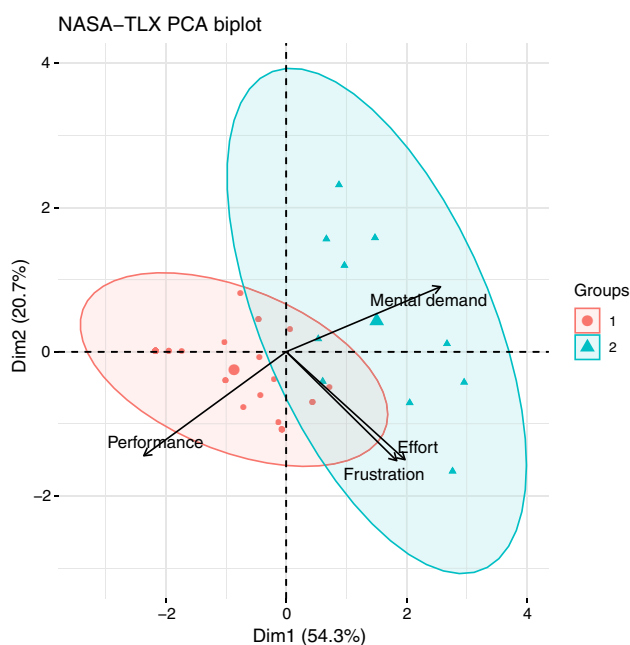


Fig. 4 Variable correlation plot for NASA-TLX items

between the perceived ease of use and the belief that people can learn to use the application very quickly, but the short length of both the arrows reveals that the two items are not well represented in the space of the first two dimensions. Also SUS items 6 and 8 are represented by close arrows, but in this case the former has a shorter arrow than the latter: this suggests that item 6 is not effectively represented by the first two components, so it is not possible to hypothesize a correlation between the perception of any inconsistencies and the idea that the system is cumbersome. The arrows of SUS items 2 and 10 are also quite close: since their length suggests that the two items are well represented, it is possible to assume a correlation between the perceived complexity and the need for an extensive training.

The bar plot in Fig. 7 depicts the square cosine (\cos^2) of SUS items on the first two components considered in Fig. 6: SUS items 2 and 10 are best represented by the two components, while SUS items 5, 1, 7 and 3 are only weakly represented, as also suggested by the shorter arrows in the diagram of Fig. 6.

Table 6 shows that the first four components account for the 77% of the total variability.

The loadings in Table 7 were obtained by performing *varimax* and *promax* rotations on the first four components. The first component is a combination of SUS items 2 and 10, which represent the perceived complexity and the need for an extensive training. The second component is a combination of SUS items 3 and 9, which represent the perceived ease of use and confidence with the application. The third component is a combination of SUS items 1 and

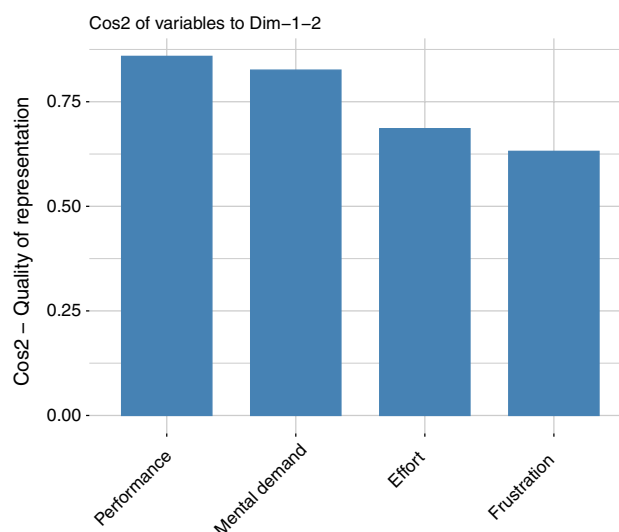


Fig. 5 Quality of representation of NASA-TLX items

Table 3 Principal component analysis of NASA items

	PC1	PC2	PC3	PC4
Standard deviation	1.4740396	0.9102530	0.8564007	0.51499947
Proportion of variance	0.5431982	0.2071402	0.1833555	0.06630611
Cumulative proportion	0.5431982	0.7503384	0.9336939	1.00000000
Eigenvalue	2.1727928	0.8285606	0.7334221	0.2652245

Cumulative proportion values that account for significant parts of the total variability are highlighted in bold

7, which represent the predisposition to use the application frequently and the thought that many people could learn to use it quickly. The fourth component depends only on SUS item 5, which indicates whether the components of the system are perceived as well integrated.

7.1.3 PCA of UEQ factors

The plot in Fig. 8 represents the relationships among UEQ factors in a space made up of the first two dimensions (which account for the 90% of the total variability): the direction of the arrows shows 3 pairs of variables *Efficiency-Perspicuity*, *Attractiveness-Stimulation* and *Dependability-Novelty* which have very similar orientation along the first two main components.

The bar plot in Fig. 9 depicts the square cosine (\cos^2) of UEQ factors on the first two components, showing that all the six UEQ factors are fairly well represented by the two components: *Stimulation* and *Novelty* are the most influential factors for variability, since they are correlated with the first

Table 4 Loadings obtained through *varimax* and *promax* rotations on the first two principal components for NASA items

	PC1	PC2
Mental demand	0.663	
	0.660	
Performance	-0.748	
	-0.751	
Effort		-0.708
		-0.709
Frustration		-0.693
		-0.694

The loadings considered significant, highlighted in bold, are those with an absolute value greater than 0.4

Table 5 Loadings obtained through *varimax* and *promax* rotations on the first three principal components for NASA items

	PC1	PC2	PC3
Mental demand	0.659	-0.243	0.118
	0.638	-0.246	0.115
Performance	-0.752	-0.221	0.109
	-0.770	-0.213	0.100
Effort		-0.944	
		-0.948	
Frustration			-0.987
			-0.989

The loadings considered significant, highlighted in bold, are those with an absolute value greater than 0.4

two principal components, whereas *Perspicuity* has the lowest influence, although the differences in the contributions of the variables are not very pronounced.

Table 8 shows that the first three and the first four components account for the 94% and the 97% of the total variability. The loadings in Table 9 were obtained by considering only the first three components: they do not reveal a clear dominance on a single component of *Attractiveness*, which has low loadings in all the three components, and *Stimulation*, whose loadings have almost the same magnitude and opposite sign on PC1 and PC2.

On the contrary, *varimax* and *promax* rotations applied on the first four components produced the loadings in Table 10, where for each factor there is a clear dominance on a single principal component. The first component can be expressed as a combination of *Attractiveness*, *Stimulation* and *Novelty*. On the other hand, *Efficiency*, *Perspicuity* and *Dependability* correspond to the second, the third and the fourth principal component, respectively.

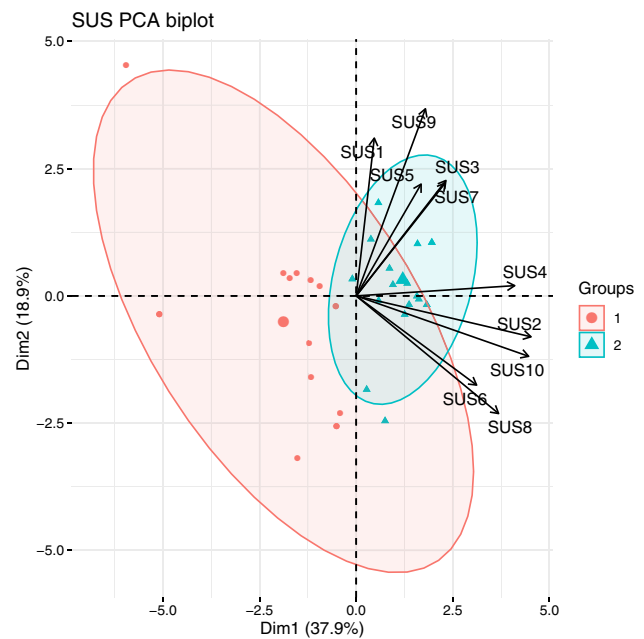


Fig. 6 Variable correlation plot for SUS items

7.1.4 PCA of all the questionnaires' factors

The plot in Fig. 10 represents the relationships among questionnaires' factors in a space made up of the first two dimensions (which account for the 76.8% of the total variability): the opposite direction of *Frustration* is negatively correlated with many other variables, such as *Perspicuity*, *Efficiency*, *Attractiveness* and *Stimulation*. Two very close arrows pointing in the same direction suggest a positive correlation between *Attractiveness* and *Stimulation*.

The bar plot in Fig. 11 depicts the square cosine (*cos*²) of all the questionnaires' items: the variables with the weakest representation are *Effort* and *Performance*; the former had a low *cos*² value also in the bar plot in Fig. 5 focusing on all the NASA-TLX items, while the latter goes from the highest score in the chart in Fig. 5 to the lowest score in the bar plot in Fig. 11; the other variables have very high *cos*² values, above 0.75, except for *Efficiency*, which however has a much higher value than *Effort* and *Performance*.

Table 11 shows that the first eight components account for the 97% of the total variability.

Varimax and *promax* rotations applied on the first four components produced the loadings in Table 12. *Attractiveness*, *Stimulation* and *Novelty*, which formed the first principal component in Table 10, now form the seventh component: therefore, they lose their importance in this wider context which includes more variables. Now the first component is positively correlated with *Frustration* and negatively correlated with *Dependability*. However, the dominance of

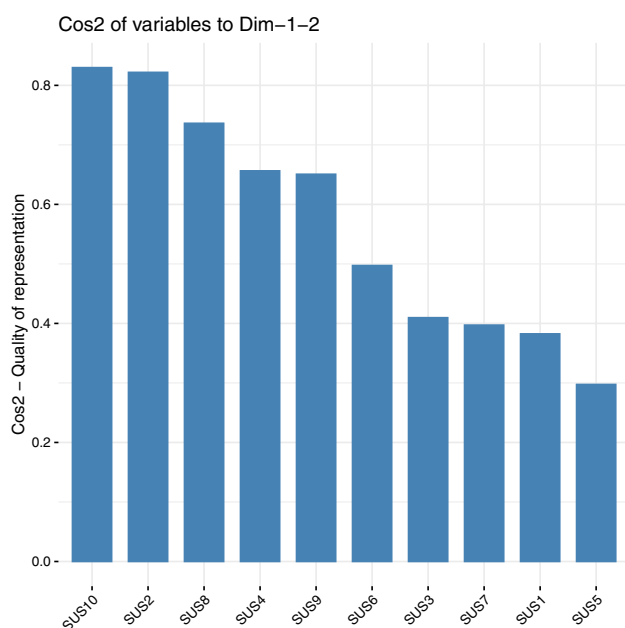


Fig. 7 Quality of representation of SUS items

Frustration on the first component is not very clear, especially according to the *promax* rotation, in which the loadings have opposite signs but not too different magnitudes on the first and last main components. *Mental demand*, *Performance*, *Effort*, *Efficiency* and *Learnability* are represented individually by components 2, 3, 4, 5 and 8 respectively. The sixth component is positively correlated with *Usability* and *Perspicuity*.

7.2 Hierarchical cluster analysis

The hierarchical procedure implemented by ICLUST performs multiple clustering steps between similar items or clusters formed in a previous stage until a stop criterion

based on *alpha* and *beta* coefficient is satisfied. An item is added to a cluster only if it can improve the cluster's internal consistency and factorial homogeneity. In the following subsections, Cluster Analysis was first applied to the items of each individual questionnaire, then to the factors describing the user experience in the UEQ questionnaire and finally to the factors of all three questionnaires taken together. In particular, a more detailed analysis is conducted at the end to identify the relationships between *Perspicuity*, *Frustration* and the individual items of the SUS questionnaire that make up *Usability* and *Learnability*.

7.2.1 Cluster analysis of NASA-TLX items

The diagram in Fig. 12 represents the clusters between items of the NASA-TLX questionnaire. Each oval connects two items forming a cluster and reports the Cronbach's *alpha* and the Revelle's *beta* reliability coefficients. *Performance* and *Mental demand*, which form up cluster C1, are again negatively correlated as suggested by PCA. *Effort* and *Frustration*, which form up cluster C2, are positively correlated, although more weakly and with significantly lower reliability coefficients. At a higher level, clusters C1 and C2 form cluster C3.

7.2.2 Cluster analysis of SUS items

The diagram in Fig. 13 represents a cluster analysis of the ten SUS items: a model based on three clusters has been adopted, since it maximizes cluster fit and pattern fit and minimizes the RMSR.

Items 4 and 10, which form up the SUS *Learnability* factor, are not directly correlated, but they belong to the same macro-cluster in which, however, other items also join. In particular, there is a significant correlation between complexity (SUS2) and the need to learn a lot of things (SUS10), which form cluster C1. The perception of the

Table 6 Principal component analysis of SUS items

	Standard deviation	Proportion of variance	Cumulative proportion	Eigenvalue
PC1	1.9471841	0.3791526	0.3791526	3.79152589
PC2	1.3739035	0.1887611	0.5679137	1.88761076
PC3	1.0852907	0.1177856	0.6856993	1.17785597
PC4	0.91362198	0.08347051	0.76916977	0.83470513
PC5	0.89384651	0.07989616	0.84906593	0.79896157
PC6	0.80294934	0.06447277	0.91353870	0.64472765
PC7	0.63373658	0.04016221	0.95370090	0.40162205
PC8	0.50136052	0.02513624	0.97883714	0.25136237
PC9	0.34943990	0.01221082	0.99104796	0.12210825
PC10	0.299199536	0.008952036	1.00000000	0.08952036

The cumulative proportion value that accounts for a significant part of the total variability is highlighted in bold

Table 7 Loadings obtained through *varimax* and *promax* rotations on the first four principal components for SUS items

	PC1	PC2	PC3	PC4
SUS1			0.641	
			0.653	-0.166
SUS2	0.485		0.107	
	0.486		0.140	
SUS3		0.587		0.114
		0.586		0.137
SUS4	0.301	0.403	-0.196	
	0.299	0.373	-0.172	
SUS5			0.130	-0.880
			0.173	-0.907
SUS6	0.438	-0.117	0.170	0.355
	0.448	-0.142	0.184	0.312
SUS7	0.172		0.625	
	0.164		0.647	
SUS8	0.446		-0.310	-0.261
	0.444	-0.120	-0.277	-0.239
SUS9	-0.116	0.685		
	-0.124	0.700		
SUS10	0.486			
	0.487			

The loadings considered significant, highlighted in bold, are those with an absolute value greater than 0.4

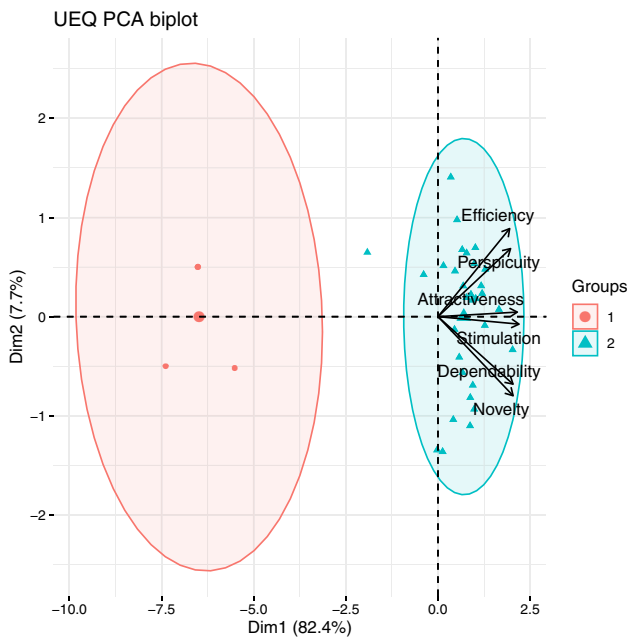


Fig. 8 Variable correlation plot for UEQ factors

system as cumbersome (SUS8) forms together with cluster C1 a higher-level cluster C3. Similarly, the need for technical support (SUS4) and the perceived inconsistency (SUS6)

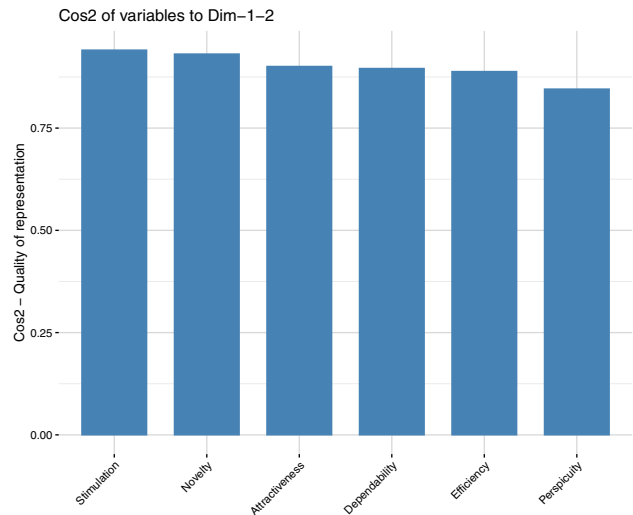


Fig. 9 Quality of representation of UEQ variables

join at higher levels to form clusters C5 and C7 respectively. However, cluster C7 presents a high discrepancy between $\alpha = 0.88$ and $\beta = 0.69$, which makes the inference on the perceived inconsistency less reliable.

7.2.3 Cluster analysis of UEQ items

The diagrams in Figs. 14 and 15 represent a cluster analysis of the 26 UEQ items, where a model based on six clusters has been adopted. In particular, there is a big macro-cluster, labelled as C19 in Fig. 14, that groups most of the items and presents the highest correlation coefficients: *valuable/inferior* and *good/bad* items, belonging to cluster C1, can be considered almost coinciding (due to correlation coefficients equal to 1), but most other coefficients are still greater than 0.9.

Other smaller clusters are represented in Fig. 15. Two items, namely *impractical/practical* and *unpredictable/predictable*, are completely isolated and do not form any clusters. The small isolated cluster C12 groups *easy to learn/difficult to learn* and *complicated/easy* items: this is in agreement with the light correlation between *Usability* and *Learnability* SUS factors suggested by Borsci et al. (2009). The medium-sized cluster C20 is the root of a hierarchical tree grouping *annoying/enjoyable*, *conservative/innovative*, *boring/exciting*, *not interesting/interesting* and *not understandable/understandable*: in particular innovative aspects makes the application enjoyable; moreover, the more interesting and enjoyable the application, the easier it is to understand.

Table 13 compares the UEQ item groupings in the three macro-clusters with the six-factor groupings in the UEQ questionnaire. The biggest cluster C19 covers almost entirely *Efficiency* and *Dependability*, half of the *Stimulation* and *Novelty* items and only one *Perspicuity* item. However, the

Table 8 Principal component analysis of UEQ factors

	Standard deviation	Proportion of variance	Cumulative proportion	Eigenvalue
PC1	2.2231187	0.8237094	0.8237094	4.94225663
PC2	0.67820729	0.07666085	0.90037029	0.45996513
PC3	0.50056497	0.04176088	0.94213117	0.25056529
PC4	0.41354861	0.02850374	0.97063492	0.17102245
PC5	0.33557104	0.01876799	0.98940290	0.11260792
PC6	0.2521559	0.0105971	1.0000000	0.06358259

Cumulative proportion values that account for significant parts of the total variability are highlighted in bold

Table 9 Loadings of the first three principal components for UEQ factors

	PC1	PC2	PC3
Attractiveness	0.269	-0.269	-0.195
	0.207	-0.264	-0.184
Perspicuity			-0.969
			-0.978
Efficiency	0.906	0.172	
	0.942	0.173	
Dependability		-0.601	
	-0.165	-0.610	
Stimulation	0.323	-0.330	
	0.261	-0.334	
Novelty		-0.653	0.111
		-0.672	0.141

The loadings considered significant, highlighted in bold, are those with an absolute value greater than 0.4

Table 10 Loadings of the first four principal components for UEQ factors

	PC1	PC2	PC3	PC4
Attractiveness	0.601		-0.155	0.252
	0.585		-0.156	0.235
Perspicuity			-0.971	
			-0.976	
Efficiency		0.957		
		0.964		
Dependability				-0.952
				-0.957
Stimulation	0.450	0.180		
	0.421	0.203		
Novelty	0.658	-0.223	0.177	-0.157
	0.694	-0.185	0.157	-0.158

The loadings considered significant, highlighted in bold, are those with an absolute value greater than 0.4

high discrepancy between $\alpha = 0.98$ and $\beta = 0.81$ in C19 makes the clustering between C17 (the subcluster that hierarchically groups the remaining items of C19) and the *attractive/unattractive* item (the most important for the *Attractiveness* factor) not very reliable.

Cluster C20 includes half of the *Stimulation* items and one item from each of the *Attractiveness*, *Perspicuity* and *Novelty* factors, even though the high discrepancy between $\alpha = 0.87$ and $\beta = 0.76$ in C20 makes the clustering between C16 and the *not understandable/understandable* item (belonging to the *Perspicuity* factor) not very reliable. The smallest cluster C12 covers only half of the *Perspicuity* items.

7.2.4 Cluster analysis of UEQ factors

The diagram in Fig. 16 represents a cluster analysis of the six UEQ factors, which shows a three-level hierarchical clustering in the form of a tree rooted in cluster C5: at the lowest hierarchy level, the closest correlation is between *Attractiveness* and *Stimulation*, followed by the one between *Dependability* and *Novelty* and then by the one between *Perspicuity* and *Efficiency*.

7.2.5 Cluster analysis of all the questionnaires' factors

The cluster diagram in Fig. 17 put together the factors from NASA-TLX, SUS and UEQ questionnaires to check whether there is any influence between factors from different questionnaires.

Attractiveness and *Stimulation*, as well as *Dependability* and *Novelty*, are still highly correlated just as in the diagram involving only UEQ factors shown in Fig. 16. They form the same hierarchical structure that was already present in the previous diagram.

Efficiency is no longer linked to *Perspicuity*, as in the diagram in Fig. 16: while in Fig. 16 it was the *Efficiency-Perspicuity* cluster that was connected to the subtree formed by *Attractiveness*, *Stimulation*, *Dependability* and *Novelty*, in Fig. 17 only *Efficiency* is connected to it.

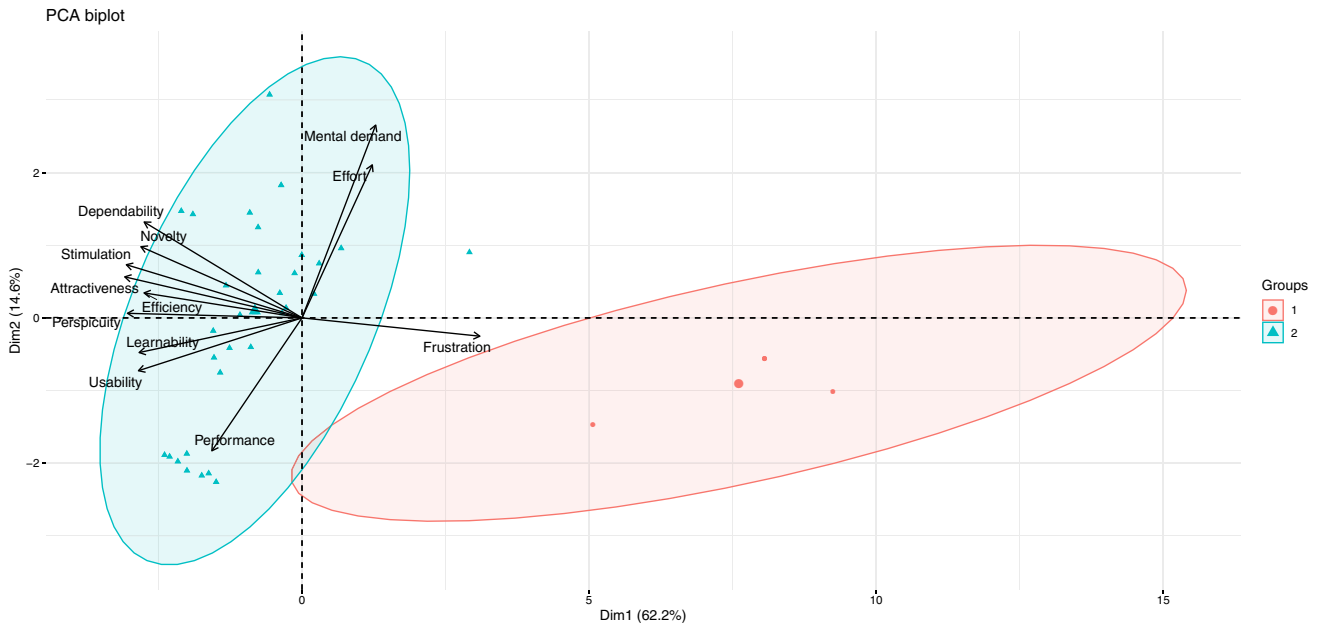
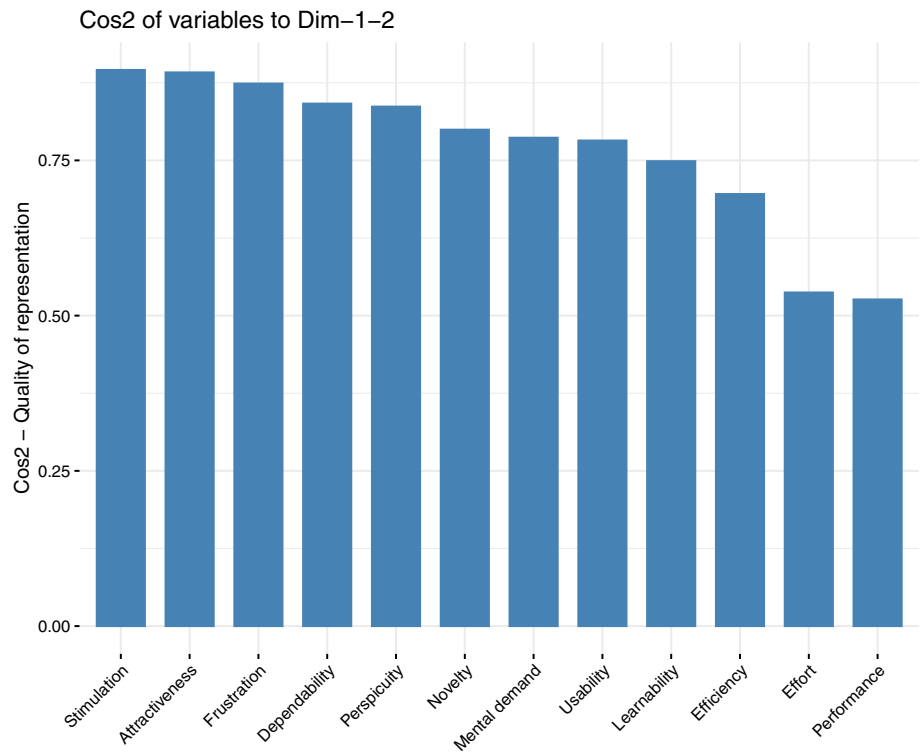


Fig. 10 Variable correlation plot for questionnaires' factors

Fig. 11 Quality of representation of questionnaires' variables



Perspicuity and *Learnability*, which refer to very close concepts, are directly grouped in a cluster showing a high correlation. *Usability*, which joins at a higher-level cluster, can be seen as a consequence of these factors.

Effort does not form any clusters with other factors: in Fig. 12 it was weakly correlated with *Frustration*, which now has a much stronger correlation with cluster C1.

The correlation between *Usability* and cluster C5 in the diagram of Fig. 17 reflects the weak connection between

Table 11 Principal component analysis of questionnaires' factors

	Standard deviation	Proportion of variance	Cumulative proportion	Eigenvalue
PC1	2.732351300	0.622145300	0.622145300	7.46574372
PC2	1.322683200	0.145790900	0.767936200	1.74949078
PC3	0.996394400	0.082733480	0.850669690	0.99280181
PC4	0.704407450	0.041349150	0.892018850	0.49618985
PC5	0.605522920	0.030554830	0.922573680	0.36665800
PC6	0.555762820	0.025739360	0.948313040	0.30887231
PC7	0.382656370	0.012202160	0.960515200	0.14642590
PC8	0.376263390	0.011797840	0.972313040	0.14157414
PC9	0.345609550	0.009953830	0.982266870	0.11944596
PC10	0.325504739	0.008829445	0.991096317	0.10595334
PC11	0.260292346	0.005646009	0.996742326	0.06775211
PC12	0.197717196	0.003257674	1.000000000	0.03909209

The cumulative proportion value that accounts for a significant part of the total variability is highlighted in bold

Table 12 Loadings of the first eight principal components for questionnaires' factors

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Mental demand		0.972						
		0.986						
Performance			-0.963					
			-0.978			0.102		
Effort				-0.978				
				-0.991				
Frustration	0.414					-0.207		-0.244
	0.363					-0.180		-0.252
Usability		0.119	-0.161	0.115		0.784		-0.124
		0.110	-0.139	0.107		0.857		-0.136
Learnability								0.921
						-0.111		0.935
Attractiveness	0.188						0.582	0.113
	0.212				-0.108		0.562	0.114
Perspicuity		-0.125	0.199	-0.117		0.540		0.123
		-0.114	0.194	-0.105		0.473		0.117
Efficiency					-0.945			
					-0.963			
Dependability	-0.873							-0.114
	-0.898							
Stimulation		-0.133			-0.152	0.111	0.461	-0.187
		-0.129			-0.158		0.413	-0.186
Novelty	-0.126				0.222	-0.162	0.669	
	-0.115				0.196	-0.118	0.718	

The loadings considered significant, highlighted in bold, are those with an absolute value greater than 0.4

Learnability and *Usability*, which was previously highlighted in literature (Borsci et al. 2009): it shows that the negative influence of *Frustration* contributes to determine *Usability* in addition to *Learnability* and *Perspicuity*, represented by cluster C1.

Mental demand and *Performance* are still correlated just as in the diagram involving only NASA items shown in Fig. 12.

Fig. 12 Cluster analysis of NASA-TLX items (Cluster fit = 0.72, Pattern fit = 0.98, RMSR = 0.09)

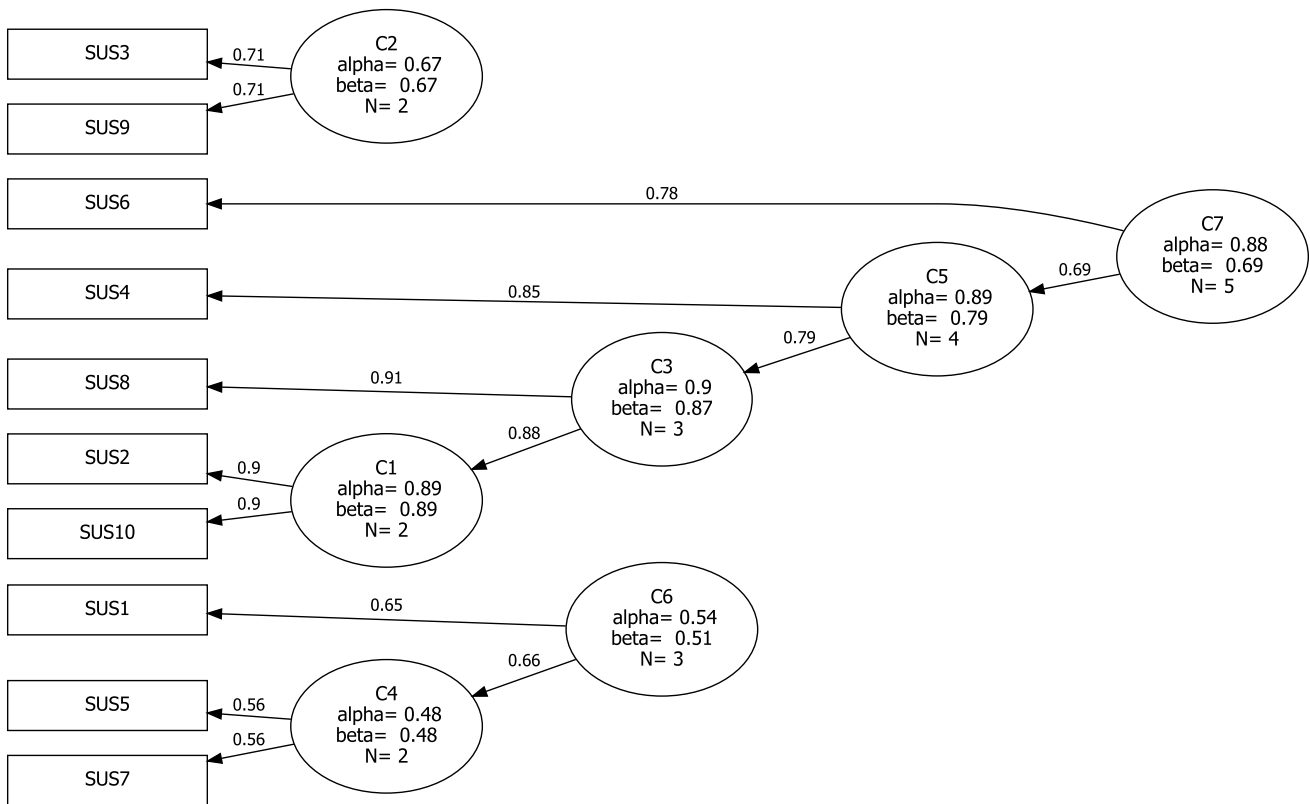
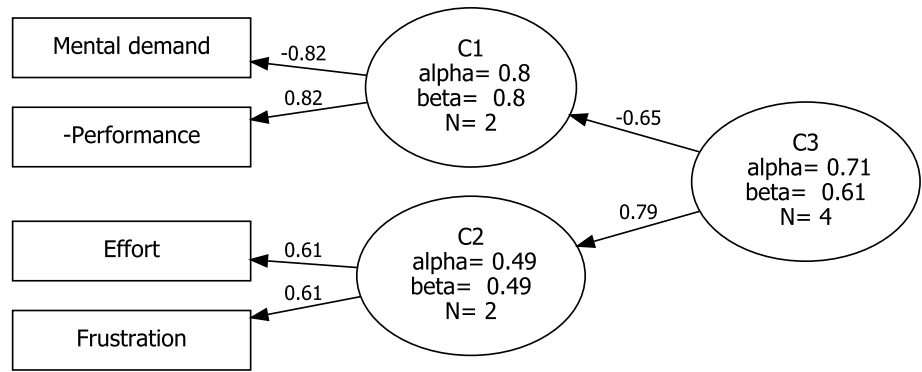


Fig. 13 Cluster analysis of SUS items (Cluster fit = 0.8, Pattern fit = 0.98, RMSR = 0.07)

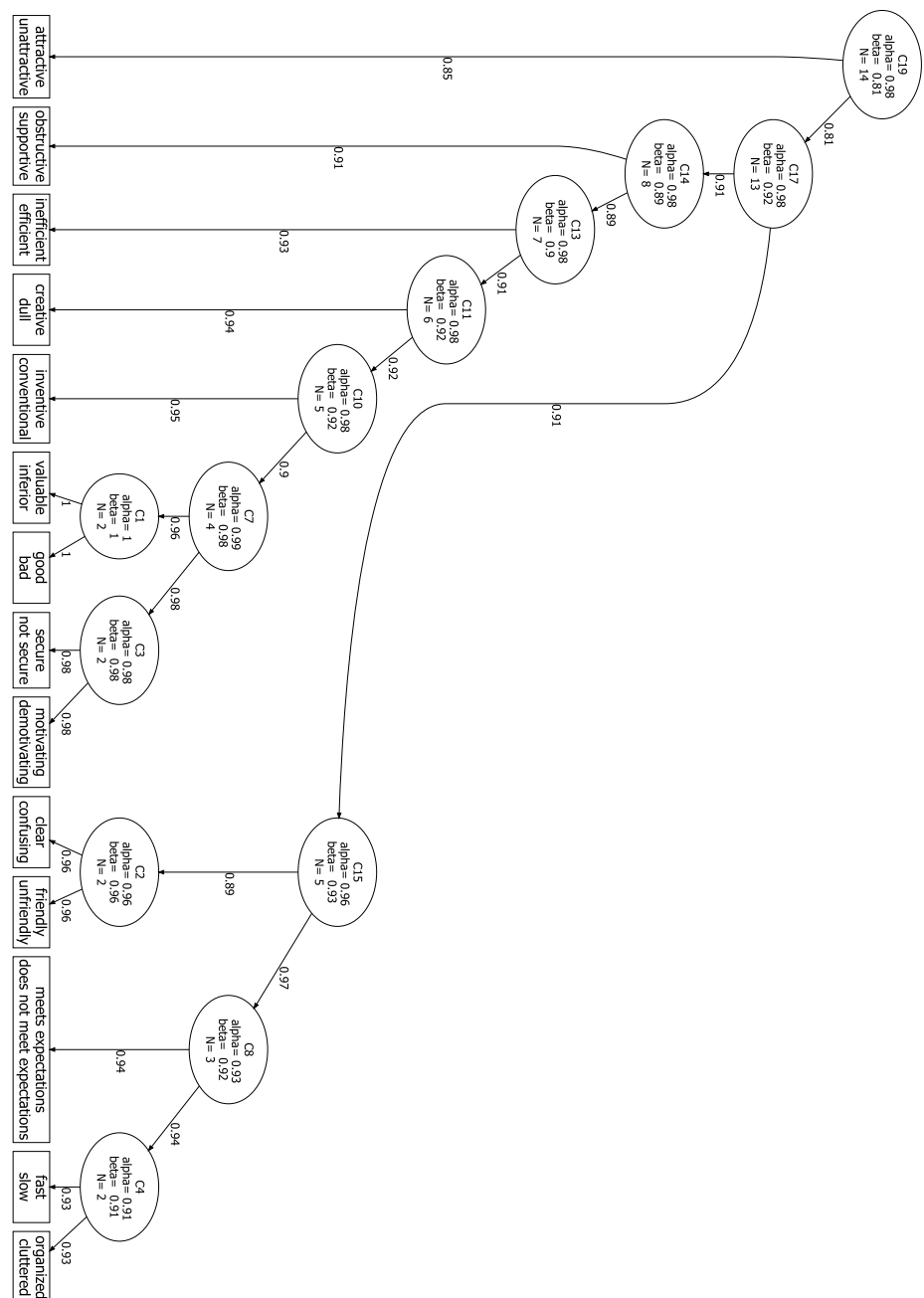
7.2.6 Usability, learnability, perspicuity and frustration

In order to understand in more detail what are the linking points between *Usability*, *Learnability*, *Perspicuity* and *Frustration*, a cluster analysis was carried out between the SUS items, the UEQ items that make up the *Perspicuity* component and the NASA-TLX item that represents *Frustration* (Figs. 18 and 19). *Frustration* is closely related to the complexity of the application (SUS2). Moreover, both these aspects increase the difficulty in learning how to use the application, as highlighted by the correlation between cluster C2 and the *easy to learn/difficult to learn* item, belonging

to the UEQ *Perspicuity* factor. In turn, these three items increase the feeling of having to learn a lot before being able to use the application effectively (SUS10), as highlighted by the correlation between cluster C4 and SUS10 item. All these aspects influence the perceived clarity in the use of the application, as highlighted by the correlation between cluster C5 and the *clear/confusing* item of the UEQ *Perspicuity* factor. The high discrepancy between α and β values for clusters C8 and C11 makes the hypothesis of a correlation with the SUS8 and SUS6 *Usability* items less reliable.

The other two items belonging to the UEQ *Perspicuity* factor represent the perceived level of complication and

Fig. 14 Cluster analysis of UEQ items (Cluster fit = 0.96, Pattern fit = 0.99, RMSR = 0.05)



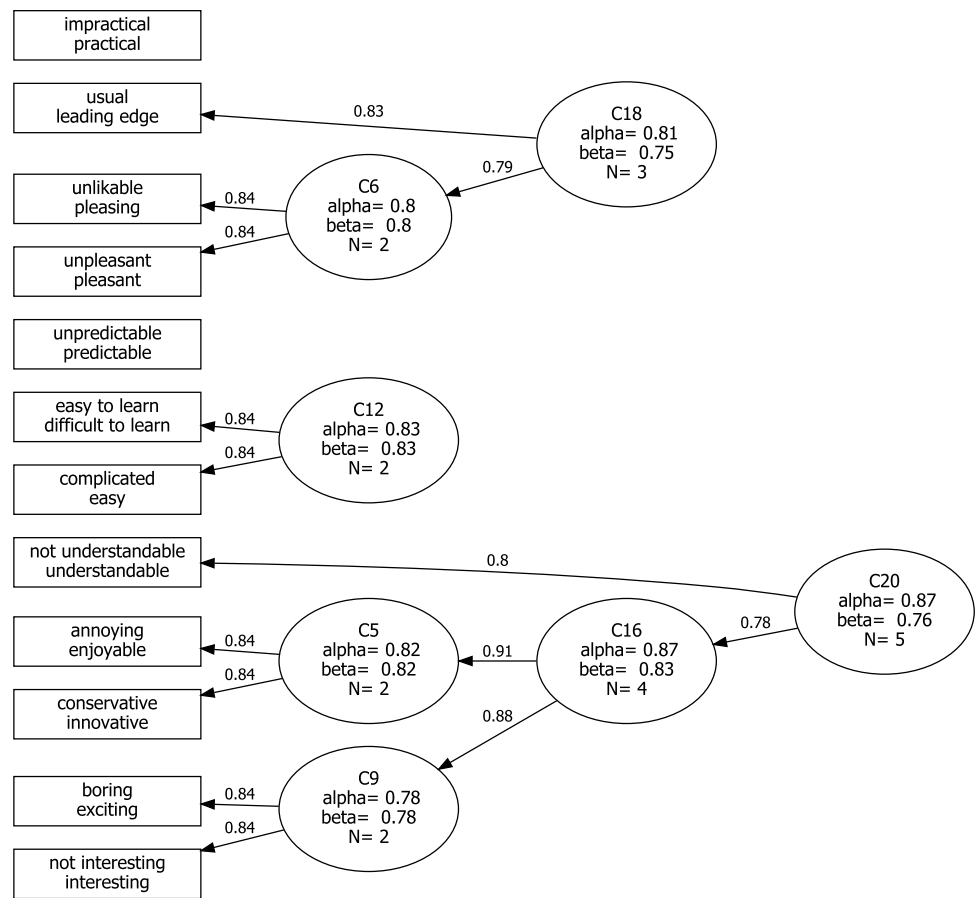
the level of comprehensibility: the former correlates with SUS4 item of *Learnability*, which concerns the need for support from an expert, while the latter correlates with SUS3 and SUS9 items of *Usability*, which concern the ease of use and the level of confidence acquired in using the application.

7.2.7 The influence of storytelling on mental workload

In order to assess in detail possible influences of storytelling on the workload, a cluster analysis was conducted between the *Mental demand* item of NASA-TLX and the items of the UEQ

questionnaire. Among the factors describing the user experience, *Dependability* (i.e. the perception of controlling the interaction) is the only one with an item, concerning unpredictability, which forms a cluster with *Mental demand* and thus seems to have influence on it (Fig. 20). *Mental demand*, on the other hand, did not correlate with any of the items of the SUS questionnaire, nor with the items of the other UEQ factors (*Attractiveness*, *Stimulation*, *Perspicuity*, *Efficiency*, *Novelty*). This suggests that mental load is not affected by usability issues, but only by certain storytelling evolutions that are considered unpredictable by some users, probably caused at least in part by the exploratory and nonlinear nature of the AR-based experience (Shilkrot et al. 2014).

Fig. 15 Cluster analysis of UEQ items (Cluster fit = 0.96, Pattern fit = 0.99, RMSR = 0.05)



8 Discussion

The connection between the worlds of usability and user experience has long been debated in the literature. In this sense, it is interesting to evaluate in the present study the connection between the SUS and UEQ questionnaires, which can be found mainly in the strong correlation between *Perspicuity* and *Learnability*, two very close concepts: the former, which refers to the ease of becoming familiar with the application, represents the more practical aspects of the latter, which refers to the ease of learning to use the application (in the sense of “understanding how to use it”).

Learnability can be further decomposed into two sub-components, represented by SUS4 and SUS10 items: the former concerns the amount of things to learn in order to use the application, while the latter refers to the practice needed under the guidance of an expert to master and apply them correctly. These two sub-components are the basis of a double connection between *Learnability* and

Perspicuity: according to the diagrams in Figs. 18 and 19, users think that any difficulties in learning to use the application can be addressed with the help of an expert, whereas the many things to learn would make the system appear complicated.

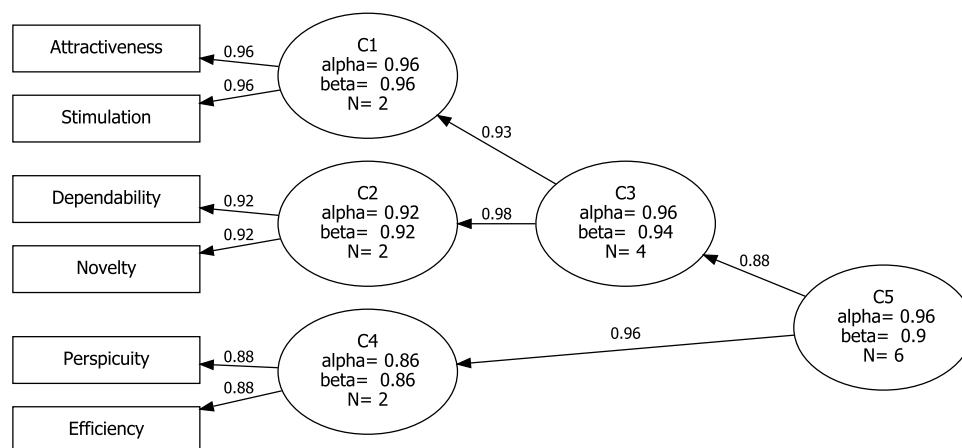
A noteworthy aspect is the absence of correlation between *Mental demand* and *Usability/Learnability*, which suggests that the user’s mind is more focused on the content of the application than on its use. The absence of correlation of *Effort* with any other factor also seems to support this hypothesis.

The fact that *Frustration* (which has an average value of only 0.561) correlates more closely to *Learnability* suggests that any difficulties were perceived mainly in the first approach to the application, as also suggested by the average *Learnability* value (3.134) which is slightly lower than the average *Usability* value (3.479).

The main factors determining the *Attractiveness* of the application are the eye-catching design (the *Novelty* factor) and the user’s perception of being able to control the

Table 13 UEQ items grouped by factors

UEQ factor	UEQ items	Cluster
Attractiveness	Annoying/enjoyable	C20
	Good/bad	C19
	Unlikable/pleasing	C18
	Unpleasant/pleasant	C18
	Attractive/unattractive	C19 (not reliable)
	Friendly/unfriendly	C19
Perspicuity	Not understandable/understandable	C20 (not reliable)
	Easy to learn/difficult to learn	C12
	Complicated/easy	C12
	Clear/confusing	C19
Efficiency	Fast/slow	C19
	Inefficient/efficient	C19
	Impractical/practical	–
	Organized/cluttered	C19
	Unpredictable/predictable	–
Dependability	Obstructive/supportive	C19
	Secure/not secure	C19
	Meets expectations/does not meet expectations	C19
	Valuable/inferior	C19
Stimulation	Boring/exciting	C20
	Not interesting/interesting	C20
	Motivating/demotivating	C19
	Creative/dull	C19
Novelty	Inventive/conventional	C19
	Usual/leading edge	C18
	Conservative/innovative	C20

Fig. 16 Cluster analysis of UEQ factors (Cluster fit = 0.99, Pattern fit = 1, RMSR = 0.04)

interaction (the *Dependability* factor), two aspects that appear to be interrelated. The latter aspect is not to be confused with *Usability*, which is unrelated to it, but should rather be understood as the way the storytelling is structured, that is, the way the content is articulated through the sequence of hot spots.

However, it should be emphasized that in the considered scenario the user is given a certain amount of freedom, as he or she can choose how long to stay at each hot spot to enjoy the associated AR content. This most likely contributes to a more pleasant experience and lightens the mental load. On the other hand, the dualism between the narrative

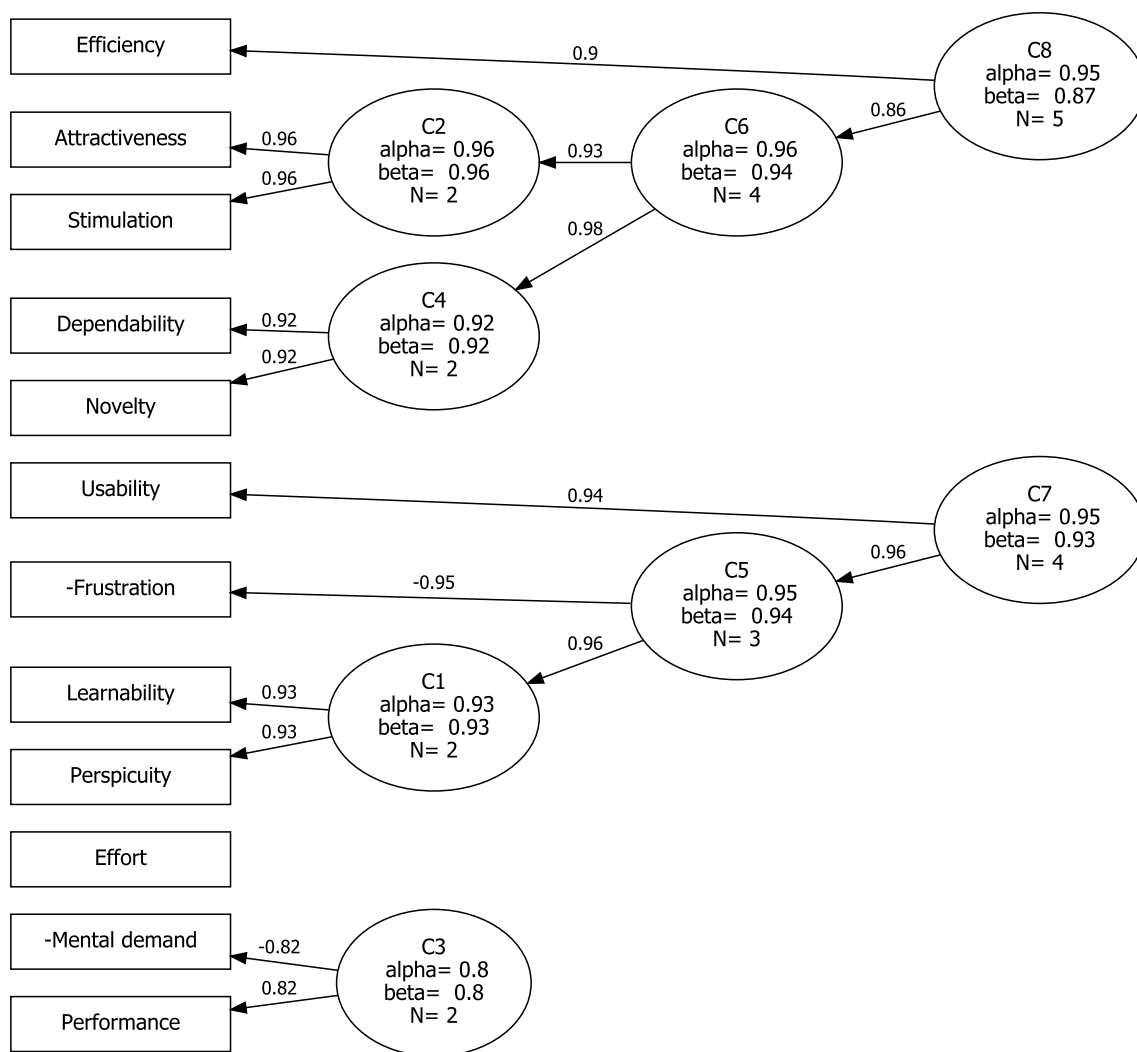


Fig. 17 Cluster analysis of all the questionnaire factors (Cluster fit = 0.97, Pattern fit = 0.99, RMSR = 0.06)

order of the story and the nonlinear exploration of the real environment could produce cognitive disorientation that could become an additional mental burden.

Even though the storytelling of the application is linked to a precise path that the user has to follow inside the Basilica, no time constraints are imposed (which is why the NASA-TLX *Temporal demand* item was not included) nor are there any particular objectives to be reached during the route. This makes the application suitable for an informal learning scenario (Lin et al. 2012), where intrinsic motivations, such as personal curiosity, are predominant and study activities are carried out as hobbies.

It would be interesting to assess how the relationships between the various factors considered in this study would change in the presence of objectives to be achieved or quizzes to be answered during the experience.

Also the inclusion of scores and gamification elements in the storytelling, such as a treasure hunt or other puzzles

along the way through the hot spots, could have a significant impact on interactions between the various factors: in particular, the desire to achieve certain game goals or scores could lead users to feel a greater impact of any usability problems on their mental load. In addition, the definition of time constraints within which to complete tasks could reveal relationships between *Effort* and other factors, as well as bring into play the time demand variable that was omitted in this study.

Moreover, the introduction of gamification elements (such as treasure hunts) could be a way to impose an order of visiting the various hot spots that could be congruent with the linearity of the story. A constrained workflow could mitigate the above mentioned cognitive misalignment, though the presence of quizzes and puzzles would require greater attention and concentration. More detailed tests will be conducted in future work to assess which of these two effects could prevail.

Fig. 18 Cluster analysis between *Usability*, *Learnability*, *Perspicuity* and *Frustration* items (Cluster fit = 0.91, Pattern fit = 0.99, RMSR = 0.06)

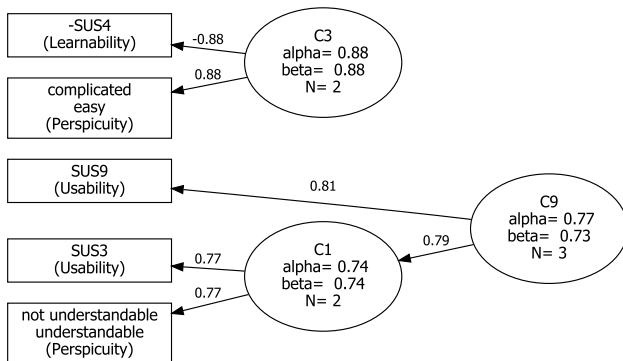
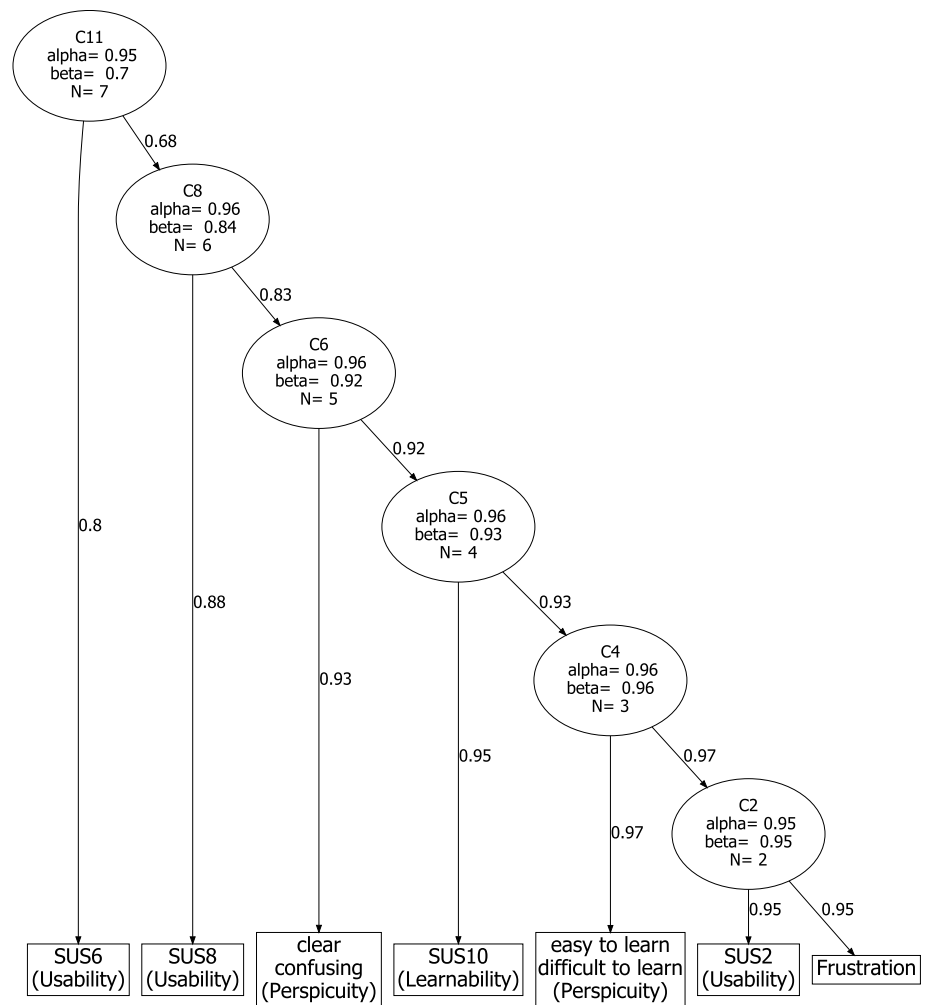


Fig. 19 Cluster analysis between *Usability*, *Learnability*, *Perspicuity* and *Frustration* items (Cluster fit = 0.91, Pattern fit = 0.99, RMSR = 0.06)

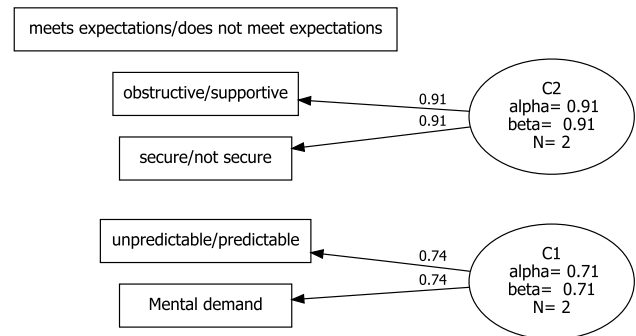


Fig. 20 Cluster analysis between the *Mental demand* item and the *Dependability* items (Cluster fit = 0.89, Pattern fit = 0.99, RMSR = 0.07)

9 Conclusions and future work

The present study analysed the relationships between usability, user experience and mental factors for a mobile application, based on AR technology, that provides the visitors of

the Basilica of St. Catherine of Alexandria with new keys to reading and thematic routes that can complement the traditional guided tour. The thematic itinerary considered in this first experimental scenario is only one of the innumerable thematic approaches that we propose to undertake, since the

Basilica with its frescoes is configured as a great book whose chapters still have much to tell to scholars and enthusiasts, a precious historical document capable of opening up to the most attentive and sensitive visitor. Each fresco can be seen as a story within history, a valuable testimony of a time gone with its customs, its creators and its enlightened patrons.

The analysis of the data collected through questionnaires after the trial showed no influence of usability or learnability on mental demand, which suggests that users were able to focus more on the content than on using the application or on learning how to use it. Moreover, usability seems to have no influence on the attractiveness of the application, which appears to be influenced more by novelty (which expresses how creative and eye-catching the design is considered) and the perception of controlling the interaction. Future work will examine whether the addition of quizzes, gamification elements and time constraints in the storytelling path will significantly change this trend.

Author Contributions LC, CG, VDL and LTDP were involved in conceptualization. VDL, CG and LTDP were responsible for methodology. CG and VDL contributed to software. LC carried out historical investigation. CG, VDL and LTDP conducted technological investigation. VDL and LTDP performed validation and took part in writing, reviewing and editing. VDL and LC curated data. VDL did formal analysis and visualization. LC, CG and VDL wrote the original draft.

Funding Open access funding provided by Università del Salento within the CRUI-CARE Agreement. The authors received no funding for conducting this study.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Ethical approval The test participants agreed with the use of the anonymously collected data for the statistical analysis presented in this research paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Assila A, Marçal De Oliveira K, Ezzedine H (2016) Standardized usability questionnaires: features and Quality Focus. *Electron J Comput Sci Inf Technol* 6(1):15–31
- Bartsch A (2012) Emotional gratification in entertainment experience. Why viewers of movies and television series find it rewarding to experience emotions. *Med Psychol* 15(3):267–302
- Bartsch A, Kalch A, Oliver M (2014) Moved to think: the role of emotional media experiences in stimulating reflective thoughts. *J Media Psychol* 26(3):125–140
- Bekele MK, Pierdicca R, Frontoni E, Malinverni ES, Gain J (2018) A survey of augmented, virtual, and mixed reality for cultural heritage. *J Comput Cult Herit* 11(2):1–36. <https://doi.org/10.1145/3145534>
- Bimber O, Raskar R (2005) Spatial augmented reality: merging real and virtual worlds. A. K. Peters; CRC Press, Natick <https://doi.org/10.1201/b10624>
- Bonacini E (2014) La realtà aumentata e le app culturali in Italia: storie da un matrimonio in mobilità / Augmented reality and cultural apps in Italy: stories on a marriage in mobility. *IL CAPITALE CULTURALE Studies on the Value of Cultural Heritage* 9:89–121. <https://doi.org/10.13138/2039-2362/740>
- Borsci S, Federici S, Lauriola M (2009) On the dimensionality of the system usability scale: a test of alternative measurement models. *Cogn Process* 10(3):193–197. <https://doi.org/10.1007/s10339-009-0268-9>
- Boskovic D, Rizvic S, Okanovic V, Slijvo S, Sinanovic N (2017) Measuring immersion and edutainment in multimedia cultural heritage applications. Institute of Electrical and Electronics Engineers Inc., vol 2017-December, pp 1–6. <https://doi.org/10.1109/ICAT.2017.8171604>
- Botrugno MC, D'Errico G, De Paolis LT (2017) Augmented reality and UAVs in archaeology: Development of a location-based AR application. In: 4th international conference augmented and virtual reality, and computer graphics (AVR 2017), Ugento, Italy, June 12–15, 2017, Lecture Notes in Computer Science (LNCS 10325), Springer, pp 261–270. https://doi.org/10.1007/978-3-319-60928-7_23
- Bozzelli G, Raia A, Ricciardi S, De Nino M, Barile N, Perrella M, Tramontano M, Pagano A, Palombini A (2019) An integrated VR/AR framework for user-centric interactive experience of cultural heritage: the ArkaeVision project. *Digital Applications in Archaeology and Cultural Heritage* 15. <https://doi.org/10.1016/j.daach.2019.e00124>
- Brooke J (1996) SUS - A quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B (eds) *Usability evaluation in industry*. CRC Press, Boca Raton
- Busselle R, Bilandzic H (2009) Measuring narrative engagement. *Media Psychol* 12(4):321–347. <https://doi.org/10.1080/15213260903287259>
- Cain B (2004) A Review of the Mental Workload Literature. NATO RTO-TR-HFM-121-Part-II
- Casciaro RDSM (2017) Basilica di Santa Caterina d'Alessandria Galatina. Congedo, Galatina
- Castaldo D (2006) Immagini della musica al tempo dei Del Balzo Orsini in CASSIANO - VETERE. Congedo, Galatina
- Cervellini F, Rossi D (2011) Comunicare emozionando. l'edutainment per la comunicazione intorno al patrimonio culturale. vol 4, pp 48–55. <https://doi.org/10.6092/issn.1828-5961/2568>
- Cisternino D, Corchia L, De Luca V, Gatto C, Liaci S, Scrivano L, Trono A, De Paolis L (2021) Augmented reality applications to support the promotion of cultural heritage. *J Comput Cult Herit*. <https://doi.org/10.1145/3460657>
- Cisternino D, Gatto C, De Paolis LT (2018) Augmented Reality for the Enhancement of Apulian Archaeological Areas. In: 5th international conference augmented and virtual reality, and computer graphics (AVR 2018), Otranto, Italy, June 24–27, 2018, Lecture Notes in Computer Science (LNCS 10850), Springer, pp 370–382. https://doi.org/10.1007/978-3-319-95282-6_27

- Cisternino D, Gatto C, D'Errico G, De Luca V, Barba MC, Paladini GI, De Paolis LT (2019) Virtual Portals for a Smart Fruition of Historical and Archaeological Contexts. In: 6th international conference augmented and virtual reality, and computer graphics (AVR 2019), Santa Maria al Bagno (Lecce), Italy, June 24–27, 2019, Lecture Notes in Computer Science (LNCS 11614), Springer, pp 264–273, https://doi.org/10.1007/978-3-030-25999-0_23
- Cooksey R, Soutar G (2006) Coefficient beta and hierarchical item clustering: An analytical procedure for establishing and displaying the dimensionality and homogeneity of summated scales. *Organ Res Methods* 9(1):78–98. <https://doi.org/10.1177/1094428105283939>
- Cronbach L (1951) Coefficient alpha and the internal structure of tests. *Psychometrika* 16(3):297–334. <https://doi.org/10.1007/BF02310555>
- Cuciniello A (2014) D'agli intendenti ammirata. La decorazione pittorica, Ortese S
- Damala A, Ruthven I, Hornecker E (2019) The MUSETECH model: a comprehensive evaluation framework for museum technology. *J Comput Cult Herit*. <https://doi.org/10.1145/3297717>
- De Paolis LT, De Luca V, D'Errico G (2018) Augmented Reality to Understand the Leonardo's Machines. In: 5th international conference augmented and virtual reality, and computer graphics (AVR 2018), Otranto, Italy, June 24–27, 2018, Lecture Notes in Computer Science (LNCS 10850), Springer, pp 320–331, https://doi.org/10.1007/978-3-319-95282-6_24
- De Paolis LT, Aloisio G, Celentano MG, Oliva L, Vecchio P (2010) A game-based 3D simulation of Otranto in the middle ages. In: 3rd international conference on advances in computer-human interactions, ACHI 2010, pp 130–133, <https://doi.org/10.1109/ACHI.2010.29>
- De Paolis LT (2013) Walking in a virtual town to understand and learning about the life in the middle ages. In: Lecture notes in computer science (LNCS 7971), Springer, pp 632–645 https://doi.org/10.1007/978-3-642-39637-3_50
- De Paolis L, Chiarello S, D'Errico G, Gatto C, Nuzzo B, Sumerano G (2021) Mobile Extended Reality for the Enhancement of an Underground Oil Mill: A Preliminary Discussion. In: De Paolis LT BP Arpaia P (ed) 8th international conference augmented and virtual reality, and computer graphics (AVR 2021), virtual event, September 7–10, 2021, Lecture Notes in Computer Science (LNCS 12980), Springer, pp 326–335, https://doi.org/10.1007/978-3-030-87595-4_24
- De Paolis LT, Aloisio G, Celentano MG, Oliva L, Vecchio P (2011a) Experiencing a town of the middle ages: an application for the edutainment in cultural heritage. In: IEEE 3rd International conference on communication software and networks, pp 169–174, <https://doi.org/10.1109/ICCSN.2011.6013802>
- De Paolis L, De Luca V, Paladini G (2019) Touchless navigation in a multimedia application: the effects perceived in an educational context. In: De Paolis LT BP (ed) 6th International Conference Augmented and Virtual Reality, and Computer Graphics (AVR 2019), Santa Maria al Bagno (Lecce), Italy, June 24–27, 2019, Lecture Notes in Computer Science (LNCS 11614), Springer, pp 348–367, https://doi.org/10.1007/978-3-030-25999-0_30
- De Paolis L, Liaci S, Sumerano G, De Luca V (2022) A video mapping performance as an innovative tool to bring to life and narrate a pictorial cycle. *Information (Switzerland)*, <https://doi.org/10.3390/info13030122>
- De Paolis LT, Aloisio G, Celentano MG, Oliva L, Vecchio P (2011b) MediaEvo project: a serious game for the edutainment. In: 3rd international conference on computer research and development, vol 4, pp 524–529, <https://doi.org/10.1109/ICCRD.2011.5763914>
- Georgiou Y, Kyza E (2017) The development and validation of the ARI questionnaire: an instrument for measuring immersion in location-based augmented reality settings. *Int J Hum Comput Stud* 98:24–37
- Hajemaesael-Gohari S, Bahaadinbeigy K (2021) The most used questionnaires for evaluating telemedicine services. *BMC Med Inf Decis Making*, <https://doi.org/10.1186/s12911-021-01407-y>
- Hall of Frescoes in the Palazzo Comunale of Tarquinia (2021) <https://www.skylabstudios.it/la-magica-sala-degli-affreschi/>
- Hendrickson A, White P (1964) Promax: a quick method for rotation to oblique simple structure. *Br J Stat Psychol* 17(1):65–70
- Herman D, Jahn M, Ryan ML (2007) Routledge encyclopedia of narrative theory. Routledge, London
- International Organisation for Standardisation (2018) Ergonomics of human-system interaction - Part 11: usability: definitions and concepts ISO 9241-11:2018. ISO 9241-11:2018(E) (November):2
- International Organisation for Standardisation (2019) ISO 9241-210: 2019 Ergonomics of human-system interaction - Part 210: human-centred design for interactive systems ergonomie. 2nd Edition. Tech. rep
- Jin Y, Ma M, Zhu Y (2022) A comparison of natural user interface and graphical user interface for narrative in HMD-based augmented reality. *Multim Tools Appl* 81(4):5795–5826
- Kaiser H (1958) The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23(3):187–200
- Kline P (2002) An easy guide to factor analysis. Routledge, London
- Laugwitz B, Held T, Schrepp M (2008) Construction and evaluation of a user experience questionnaire. In: Holzinger A (ed) HCI and Usability for education and work, lecture notes in computer science, Springer, Berlin, Heidelberg, vol 5298, pp 63–76, https://doi.org/10.1007/978-3-540-89350-9_6
- Lewis JR (2018) Measuring perceived usability: the CSUQ, SUS, and UMUX. *Int J Human-Comput Interact* 34(12):1148–1156
- Lewis JR, Sauro J (2009) The factor structure of the system usability scale. In: Kurosu M (ed) Human centered design, lecture notes in computer science, Springer, Berlin, Heidelberg, pp 94–103, https://doi.org/10.1007/978-3-642-02806-9_12
- Lin AC, Fernandez WD, Gregor S (2012) Understanding web enjoyment experiences and informal learning: a study in a museum context. *Decis Support Syst* 53(4):846–858. <https://doi.org/10.1016/j.dss.2012.05.020>
- Liu Y, Spierling U, Rau L, Dörner R (2021) Handheld vs. head-mounted AR interaction patterns for museums or guided tours. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST* 377:229–242. https://doi.org/10.1007/978-3-030-76426-5_15
- MAUA Museum (2021) <https://mauamuseum.com/>
- Miller C (2019) Digital storytelling: a creator's guide to interactive entertainment. CRC Press, Boca Raton
- Montinari M (1978) La basilica cateriniana di Galatina; con due saggi di Franco Silvestri; Acquerelli di Pietro Cavoti. Editrice Salentina, Galatina
- Moustafa K, Luz S, Longo L (2017) Assessment of mental workload: a comparison of machine learning methods and subjective assessment techniques. In: Longo, L., Leva, M. (eds) Human Mental Workload: Models and Applications. H-WORKLOAD 2017. Communications in Computer and Information Science, vol 726. Springer, Cham. https://doi.org/10.1007/978-3-319-61061-0_3
- Murray J (1997) Hamlet on the holodeck: the future of narrative in cyberspace. Free Press
- Nasa NA, Administration S (2010) NASA TLX: Task Load Index. *Planta Medica*

- Okanovic V, Ivkovic-Kihic I, Boskovic D, Mijatovic B, Prazina I, Skaljo E, Rizvic S (2022) Interaction in extended reality applications for cultural heritage. *Appl Sci* (Switzerland). <https://doi.org/10.3390/app12031241>
- Oliver M, Raney A (2011) Entertainment as pleasurable and meaningful: identifying hedonic and eudaimonic motivations for entertainment consumption. *J Commun* 61(5):984–1004. <https://doi.org/10.1111/j.1460-2466.2011.01585.x>
- PTC (2021) Vuforia Engine Developer Portal. <https://developer.vuforia.com/>
- Ping J, Liu Y, Weng D (2020) Study on Mobile AR Guide System to Enhance User Experience in Cultural Heritage Sites. *Institute of Electrical and Electronics Engineers Inc.*, pp 79–86, <https://doi.org/10.1109/ICVRV51359.2020.00027>
- Porter Abbott H (2014) *The Cambridge introduction to narrative*, 2nd Edn. Cambridge University Press, <https://doi.org/10.1017/CBO9780511816932>
- Revelle W (1978) ICLUST: a cluster analytic approach to exploratory and confirmatory scale construction. *Behav Res Methods Instrum* 10(5):739–742. <https://doi.org/10.3758/bf03205389>
- Revelle W (1979) Hierarchical cluster analysis and the internal structure of tests. *Multivar Behav Res* 14(1):57–74. https://doi.org/10.1207/s15327906mbr1401_4
- Roth C, Koenitz H (2016) Evaluating the user experience of interactive digital narrative. *Association for Computing Machinery, Inc*, pp 31–36
- Ryan ML (2006) *Avatars of story*, vol 17. University of Minnesota Press, Minneapolis
- Santos M, Polvi J, Taketomi T, Yamamoto G, Sandor C, Kato H (2015) Toward standard usability questionnaires for handheld augmented reality. *IEEE Comput Graphics Appl* 35(5):66–75
- Santos M, Polvi J, Taketomi T, Yamamoto G, Sandor C, Kato H (2014) Usability scale for handheld augmented reality. *Association for Computing Machinery*, pp 167–176
- Shilkrot R, Montfort N, Maes P (2014) nARratives of augmented worlds. *Institute of Electrical and Electronics Engineers Inc.*, pp 35–42, <https://doi.org/10.1109/ISMAR-AMH.2014.6935436>
- Terracotta Warriors of The First Emperor (November, 2021) <https://www.fi.edu/exhibit/terracotta-warriors>
- Thurstone LL (1947) *Multiple-factor analysis: a development and expansion of the vectors of mind*. University of Chicago Press
- Unity Technologies (November, 2021) *Unity*. <https://unity.com/>
- Vacca N (1955) *Il Sacramento del Matrimonio e il ritratto di Maria d'Engnien in Santa Caterina di Galatina e in S.M. di Cásule presso Copertino. "Nuptiae Sallentinae" per lo spozalizio di Fausta Vacca con Augusto Giovannini*, Lecce: Editrice Salentina
- Venkatesh V, Morris M, Davis G, Davis F (2003) User acceptance of information technology: toward a unified view. *MIS Quart Manag Inf Syst* 27(3):425–478
- Vert S, Andone D, Ternauciu A, Mihaescu V, Rotaru O, Mocofan M, Orhei C, Vasii R (2021) User evaluation of a multi-platform digital storytelling concept for cultural heritage. *Mathematics*. <https://doi.org/10.3390/math9212678>
- Vlahakis V, Ioannidis N, Karigiannis J, Tsotros M, Gounaris M, Stricker D, Gleue T, Daehne P, Almeida L (2002) *Archeoguide: an augmented reality guide for archaeolog sites*. *IEEE Comput Graphics Appl* 22(5):52–60. <https://doi.org/10.1109/MCG.2002.1028726>
- Vosinakis S, Koutsabasis P (2018) Evaluation of visual feedback techniques for virtual grasping with bare hands using Leap Motion and Oculus Rift. *Virtual Reality* 22(1):47–62
- Zoellner M, Stricker D, Bleser G, Pastarmov Y (2007) iTACITUS - Novel Interaction and Tracking Paradigms for Mobile AR. In: *The 8th international symposium on virtual reality, archaeology and cultural heritage, VAST 2007, Incorporating the 5th EURO-GRAPHICS Workshop on Graphics and Cultural Heritage*, Budapest, November 26–29, pp 110–117

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.