Multiuser Interaction with Hybrid VR and AR for Cultural Heritage Objects

Yue Li
International Doctoral Innovation Centre & NVIDIA Joint-Lab on Mixed Reality
University of Nottingham
Ningbo, China
yueli@nottingham.edu.cn

Eugene Ch'ng
NVIDIA Joint-Lab on Mixed Reality
University of Nottingham
Ningbo, China
eugene.chng@nottingham.edu.cn

Shengdan Cai
Faculty of Humanities and Social Sciences & & NVIDIA Joint-Lab on Mixed Reality
University of Nottingham
Ningbo, China
shengdan.cai@nottingham.edu.cn

Simon See
NVIDIA AI Technology Centre
NVIDIA
Singapore
ssee@nvidia.com

***Abstract*— This research investigates the factors and ways in which users initiated conversations and engage in interactions in a hybrid virtual environment using a combination of Virtual Reality (VR) and Augmented Reality (AR) devices. The research was done in the 'spirit of the ancient Silk Road' where trade brought in exchange of ideas, cultural influence and cross-border communications. The notion of a 21st century Silk Road is necessarily digital, over the Internet and based around 3D cultural heritage objects. Digi-Capital's Report forecasts the revenue of AR and VR to be US$150b by 2020. We projected that VR and AR will become pervasive, much like the Social Web and the universal ubiquity of mobile devices such as smartphones and wearables. Here, we conducted a user study exploring users' acceptance of the use of hybrid VR and AR for cultural heritage, and investigated the social nature of multiple co-located user interaction. We adapted the UTAUT questionnaire in our experiment and found that social influence has positive effects on performance expectancy and effort expectancy, which generate positive effects on user behavioural intention. This study pioneers the future design and use of hybrid VR and AR technology in cultural heritage specifically, and in other application areas generally by highlighting the significant role that social influence plays in enhancing users' behavioural intention facilitated by different immersive devices.**

Keywords—virtual reality, augmented reality, hybrid VR AR, technology acceptance, interaction design, social interaction, cultural heritage, heritage objects

I. Introduction

The ancient Silk Road was a network of multiple trade routes that played a significant role in the expansion of trade and culture between the East and the West. In particular, heritage objects associated with the past can encapsulate both history and the remarkable stories of the time. They were both objects of trade, of symbolic importance and of personal sentiments from which conversations occurred and money and friendships exchanged. In contemporary times the objects are preserved, distributed in museums all over China and other countries. While a significant number of objects are not on display, those of high importance and on display are not accessible, at least for people who do not have the means to travel to museums in multiple locations in order to view the objects. The advent of digital technology and interface devices could provide not only global accessibility, but communications, the construction of narratives and most importantly, the 'spirit of the ancient Silk Road' - social interaction and exchange of ideas, all supported by transformative digital technologies and the Social Web (Web 2.0).

Literature in museum studies has in recent times emphasised the importance of the social nature of visitors and highlighted that collective activities in visitor experience do strengthen memorability, enhance engagement and contribute to the general user experience [1]. Digital technology is at a stage where highly realistic objects and environments, real-time interactivity, and multiuser virtual experience have become possible [2], [3]. The 'in-the-wild' study of technology acceptance of Virtual Reality (VR) and Augmented Reality (AR) devices suggests that popular devices can be adopted for use fairly quickly (<3 minutes) by first-time users [4]. As far as we know, there has been no study to-date investigating technology acceptance, user behaviour and the mechanisms of interaction and communication between hybrid VR and AR devices, especially in their use for accessing cultural heritage objects. Since applications using VR and AR devices are object-centric, i.e., both technologies present heritage objects to users via displays, it brings into question how they would work in a hybrid environment when multiple users attempt to interact and communicate with different devices around virtual objects. The research explores the use of hybrid VR and AR within a multiuser application and investigates users' acceptance of the technologies, proposing and testing a theoretical model incorporating social influence, performance expectancy, effort expectancy, and behavioural intention. In our work, we created a virtual environment with digitised Chinese relics with interface technologies shared between the HTC Vive and a mobile AR app, providing users with dynamic interaction mechanism that goes beyond the exhibition space, allowing them to share their interpretation of artefacts. A questionnaire [5] was adapted to use in the study to measure technology acceptance. These quantitative data, together with qualitative data from interviews, were collected to provide a highly objective and in-depth view of users' acceptance towards the use of VR and AR in cultural heritage with a social context and the mechanisms of multiuser interactions around artefacts.

II. Background

VR and AR have in the last two years gained widespread attention for research and application with affordable headsets such as the PC-based Oculus DK2, HTC Vive, and Sony PlayStation VR, as well as some mobile VR devices including Samsung Gear and Google Cardboard. VR has a long history of research, beginning with its initial use within the military simulations and medical applications domain [6]. VR has since been applied to many other areas. In our article we focus on cultural heritage, an area with a large community actively adopting VR and AR as a means of access to sites, monuments and objects. Using VR, visitors can travel back in time and experience the past [7], such as an underwater adventure to see some of Earth's earliest inhabitants [8], or to explore a 10,000 year old Mesolithic village [9], [10], and see the world's oldest calendrical system at work [11]. While the aforementioned VR applications substitute the real world with completely simulated environments, AR augments supplementary information for the real world over a display. An example is a portable AR system designed to augment fragmentary aspects of a sculpture [12]. The use of VR and AR technology for cultural heritage demonstrates new ways of accessing information. We believe that it is likely to be more widely used for visitors, for more engaging learning experience and social interaction around cultural heritage objects and environments. Access to cultural heritage is now possible outside of cultural institutions, in more private spaces such as homes for example, as VR and AR devices have been made consumer level devices, available and affordable. New ways of accessing information are making it possible for museums to open up their collections, a solution for museums lacking in space and appropriate conditions for displaying collections. Many artefacts are preserved in warehouses due to their sensitivity to light and humidity. However, by digitising these objects and presenting them using interactive VR and AR technologies, the accessibility of objects for audiences becomes a possibility, since they are no longer confined to the limited physical spaces of museums [13]. Digitalised objects could be presented at a high fidelity, allowing users to inspect objects at close range, and form interpretations without the risk of breaking them.

While there have been various degrees of participant studies in the use of VR and AR in cultural heritage, few have considered their use within a social context, and none using hybrid devices simultaneously. The nature of AR applications, if using a mobile device, enables multiple users to share a display device and see objects together. VR being fully immersive isolates users within a virtual environment. We believe that the two devices needs to be combined in some ways.

Previous research has emphasised the importance of social interactions in museums as they tend to contribute to collaborative learning through discussions, debates which lead to deeper reflections on the subject [14]. These are important and should not be compromised when introducing emerging digital technologies. Therefore, an exploratory research in the use of hybrid VR and AR has now become important for understanding social interactions around artefacts. A hybrid VR and AR system could also be used for domains outside of cultural heritage, and contribute to knowledge in the design of multiuser, multidevice environments. We felt that user acceptance of hybrid devices is initially the key to better design. Several models of technology acceptance have been developed. A suitable model is the Technology Acceptance Model (TAM). It suggests two notable external variables that influence users' acceptance of a technology: perceived usefulness (PU) and perceived ease-of-use (PEOU) [15]. Haugstvedt and Krogstie [16] studied the acceptance of mobile AR using an adapted TAM model upon an application with information and photographs of a historic street. Aside from PU and PEOU, they have also studied perceived enjoyment, behavioural intention, and individual variables in their experiments, arguing that perceived usefulness and perceived enjoyment have a direct impact on users' intention in the use of the application. However, the model is limited as it only investigated the individual use of technology when a historical street is a public environment where interactions among pedestrians are common. Despite the significant use of the TAM model, it has been widely criticised, especially for its lack of consideration in the social process and consequences of the technology use. This leads to the refinement and expansion of the original model. Venkatesh *et al.* [5] proposed the Unified Theory of Acceptance and Use of Technology (UTAUT) identifying four key constructs: 1) Performance Expectancy (PE), 2) Effort Expectancy (EE), 3) Social Influence (SI), and 4) Facilitating Conditions (FC), taking account of the social factors. Therefore, we adapted this model in our user study. Carrozzino and Bergamasco summarised several examples of VR use in cultural context and noted that users rated an overall high acceptance for two immersive VR system [17]. However, some issues related to the VR installations were identified, this includes the diverse expertise required for its implementation, users' reluctance to adopt 'intimidating' mechanical appliances, and most importantly, the lack of interaction and limitations of supporting multiple users in immersive systems. This confirms our argument of the necessity to further investigate user acceptance of multiuser, multidevice application with VR and AR, especially the viability of hybrid systems for facilitating interaction, engagement, and communication between users in our proposed social virtual environment.

Social influence, defined as "the degree to which an individual perceives that important others believe he or she should use the new system" [5], is a significant aspect to consider when designing systems with social support. There are design principles for facilitating human-human interaction in order to satisfy their motivational requirements on social and psychological relatedness [18]. Moreover, it was suggested that a system that support social use makes users feel related and is more likely to engage users in future use [19]. Considering that our system is a multiuser, multidevice system where social interactions will occur throughout the entire experience, our hypothesis is that social influence is likely to affect users' expectancy on performance and effort. In addition, among the four key constructs identified in UTAUT, the first three were confirmed to have positive effects on behavioural intention. Therefore, we proposed the following hypotheses:

H1. Social influence has a positive effect on performance expectancy.

H2. Social influence has a positive effect on effort expectancy.

H3. Performance expectancy has a positive effect on behavioural intention.

H4. Effort expectancy has a positive effect on behavioural intention.

III. METHODOLOGY

This section describes our setup for the virtual environment and the interaction design.

A. Experimental Virtual Environment

We created a virtual environment for testing our hypotheses. The environment allows social interaction between VR and AR devices. It hosts our initial collection of cultural heritage artefacts.

A Local Area Network (LAN) was established to synchronise different users' positions and interactions with the interactable objects. Six photogrammetry reconstructed objects were arranged in a circle, placed on top of a pedestal so that each of them is equally accessible (Fig. 1). Each object has a label containing text and images. The object as an interface in itself implements interaction modes, allowing it to be picked up and manipulated using tracked hand-held controllers.

Fig. 1. Exhibition room with objects and labels

1) Overview of Environment

The virtual environment was developed in Unity (version 2018.1.0f2) built on a professional VR workstation with NVIDIA Quadro M6000 24GB graphics card and HTC Vive. Our workstation has an Intel i7 2.4GHz 12-core CPU, 64GB of RAM, and 2TB HDD. The graphics were displayed within an HMD with a 90 Hz refresh rate, a horizontal 100-degree field of view and a total resolution of 2160*1200 pixels. Two sensors were used to locate user positions with respect to the virtual environment and two hand-held controllers were used for interacting with objects. AR users made use of Samsung Galaxy S7 for accessing the virtual objects and their information. Vuforia AR SDK was used in the AR development [20]. The system was available in both English and Chinese, with additional language support available as part of the 'spirit of the ancient Silk Road'.

The environment was kept simple but consistent, focusing only on the objects as we considered objects as the interface between VR and AR devices. Photogrammetry models in their original format have extremely large polygon count. Whilst they do not affect our high-end setup, we have processed them so that they could be used in low-end systems. The models were retopologised – the geometry were recreated with a new mesh by modifying the edges and meshes. Retopology makes it easier to texture and animate an object model and could help clean up noises within data in the raw model produced from the photogrammetry process. After retopology and texture baking within Blender, the models were exported to Unity in the scene creation and development where the models were enhanced further with texture and normal mapping, thus reconstructing the representation to a visual quality asymptotic to the original object. The detailed workflow is described in Fig. 2.

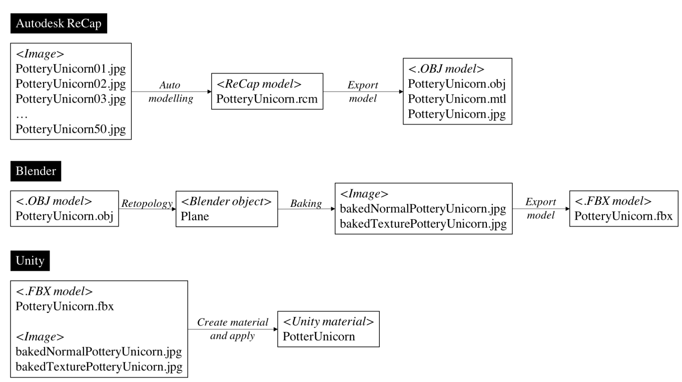
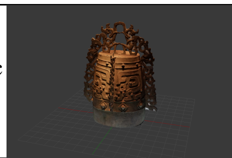
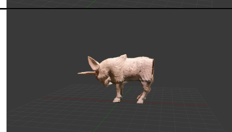
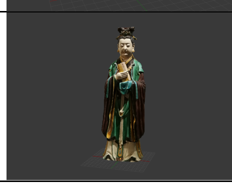


Fig. 2. Process workflow for digital cultural heritage objects, take the Pottery Unicorn for example (object #4, see Table I.)

2) Overview of Objects

The virtual environment contained a mixed collection of artefacts from different time periods and materials. Details of the objects are shown in Table I.

TABLE I. ARTEFACTS INFORMATION

#	Name	Picture	Size	Time Period	Museum
1	Bronze Music Instrument		Height: 63 cm	Western Zhou	Tianjin Museum
2	The Bronze Mask with Protruding Pupils		Height: 66 cm Width: 138 cm	Shang	Sanxingdui Museum
3	Pottery Figure of a Standing Lady		Height: 75.5cm Width: 26.6 cm	Tang	National Palace Museum, Taipei
4	Xie Zhi (Pottery Unicorn)			Northern Wei	Shaanxi History Museum
5	Tri-coloured Camel		Height: 87 cm	Tang	Nanjing Museum
6	Figure of an Assistant to the Judge of Hell		Height: 148 cm Width: 36 cm Depth: 20 cm	Ming	British Museum

3) Interaction Design

The system supported both VR and AR devices and interactions between users:

- VR users were donned with a headset with accompanying hand-held controllers.
- AR users were positioned in the real environment, using a mobile device.

a) VR Interactions

Each VR user was represented by a simple virtual avatar. The user's position and orientation relative to the virtual environment were synchronised across the network, and read from embedded sensors in the headset and controllers. Users were able to see each other's real-time movements as reflected by the avatar. The AR user had a limited view, the layout of virtual objects and the virtual avatars were not visible in AR view. However, the virtual avatar of the AR user and the object being accessed could be viewed by VR users. The virtual avatar for AR was static, the position existed in relation to the object but movements were not synchronised. The object being accessed, however, was synchronised as it is manipulated, in both position and orientation as users engaged with it. For example, when the AR user manipulated an artefact, the VR users could see the artefact being moved within the VR environment. We consider VR and AR as being from different worlds, using different devices and therefore, the need for objects to be an interface connecting them. This is the value and contribution of our present research.

Within the virtual environment, users interacted with virtual objects by picking up and putting down using hand-held controllers. Objects being picked up were highlighted with an outline, and were snapped to its original position on the pedestal when the pickup trigger was released (see Fig. 3). A virtual environment need not have glass cases as in the museums, as such users could freely access the artefacts and appreciate them from all angles. This is an important benefit of virtual reality simulations, it enhances access to artefacts, and therefore the personal appreciation and interpretation of objects. Although it was technically feasible to toss and break and object, we limited the action for the purpose of showing respect towards cultural heritage. This is again, the benefit of using VR, for destructive actions can be made impossible. Using the HTC Vive, users could move around in the virtual environment by walking around in the real world, providing them with a direct mapping between their physical movements and their positions in the virtual environment. This reduces the possibility of simulator sickness generally caused by teleportation and sudden directional change.

b) AR Interactions

Our design rationale for AR interaction begins from the scenario of AR usability. Research found that for users co-located with the VR player, even though they were not immersed in a virtual environment, they would be keen to interact with HMD users [21]. They specifically supported the idea of asymmetric approaches, such as having different views of a scene and different methods of interactions within the same scene. We believe that AR technology fits well with this scenario when a user wants to be part of a VR experience, but does not have access to a HMD. Such an experience provides different levels of immersion and a completely different display approach, providing users of both VR and AR a unique experience. Furthermore, we see that AR supported devices such as smartphones and tablets are comparatively more affordable, as such, they will have greater accessibility and more potentials for social use than VR. Another reason which AR is seen to be more useful is that motion sickness in VR can be an issue. In such cases, AR can be an ideal alternative approach for accessing cultural heritage objects.

In our study, we provided AR users with a cube with six faces (like a dice, with a side of 6 cm), each with an image pattern that a cultural heritage object was augmented on (see Fig. 4). We designed the physical cube as an embodiment of the six artefacts as well as for their interactions. This followed the principle of Embodied Interaction [22] in that it allowed users to manipulate the cultural artefacts through the engaged interactions with the cube. By rotating the physical cube, the

AR user could observe an object from different angles within the real world. Together with the artefact information, the object became an interface between VR and AR users as it created their shared experience. When an AR user looked at the augmented object, the same object was rotated triggering a sound effect within the virtual environment, providing visual and auditory cues to draw their attention towards the object. This necessarily promotes initial engagement between users with different devices.

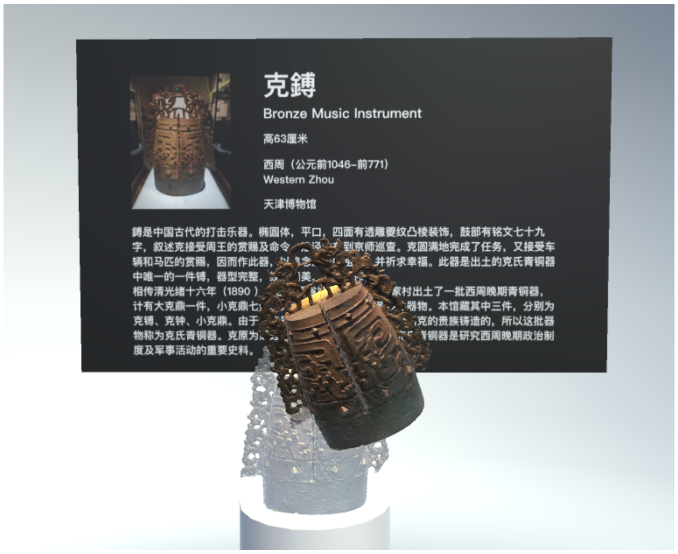


Fig. 3. The visual cue indicating that the object being picked up can be snapped back to its original location

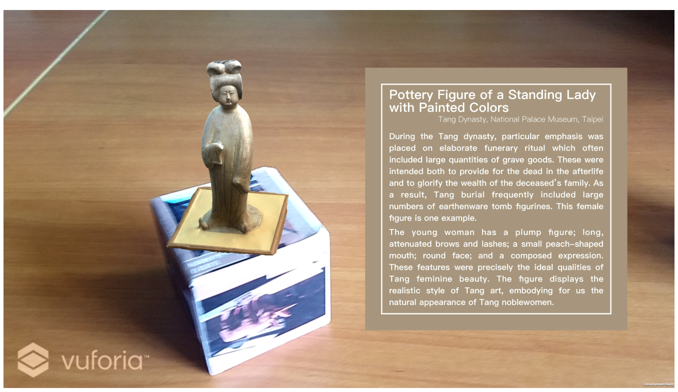


Fig. 4. The physical AR cube with augmented artefact model and information label displayed within the mobile phone's AR viewport

B. *User Study*

We conducted a user study investigating users' acceptance of the use of hybrid VR and AR for cultural heritage, as well as the social nature of multiple co-located user interaction. A pre-experiment questionnaire and an adapted UTAUT questionnaire were used to yield data from the experiments. Qualitative data were collected from our observations, and also provided by users during the interview as they reflect on their experience. Each set of experiment involves paired users – a VR user and the accompanying AR user.

1) *Demographics Group*

In total, we involved 52 users (28 males, 24 females) as 26 pairs in the study. Users were recruited randomly through the university's offline and online public channels, via emails, posters, and social media posts. The demographics of users are given in Table II. Most users were aged between 18 and 34. A majority of them were Chinese and most pairs of users were acquainted with each other, only 6 pairs of study took place with strangers.

TABLE II. DEMOGRAPHICS OF USERS

Category	Item	Frequency	Percentage
Gender	Male	28	53.85%
	Female	24	46.15%
		52	100.00%
Age	18-24	30	57.69%
	25-34	19	36.54%
	35-44	2	3.85%
	45-54	1	1.92%
		52	100.00%
Chinese	Yes	47	90.38%
	No	5	9.62%
		52	100.00%
Acquainted	Yes	20	76.92%
	No	6	23.08%
		26	100.00%

2) *Study Environment*

The experiments took place at the NVIDIA Joint-Lab on Mixed Reality, an NVIDIA Technology Centre at the University of Nottingham's China campus. A 2.5 m x 2.5 m space was used, calibrated with the HTC Vive for freedom of movement. During the experiment, ethnographic studies were conducted to observe and record user actions and the conversations without interrupting them.

C. *Data Collection and Data Analytics – Quantitative and Qualitative*

1) *Data Collection*

The pre-experiment questionnaires used in our study had three sections: user demographics, including gender, age, profession; previous experience and familiarity with 3D gaming, VR and AR; and previous experience with museum visits, interests in cultural heritage, and knowledge of Chinese artefacts.

After each experiment, users were asked to fill in a technology acceptance questionnaire, in which items were taken from the original UTAUT questionnaire and adapted to fit within the context of our hybrid VR and AR research. Our experiment was designed to test our hypotheses about the relationship between social influence, performance expectancy, effort expectancy, and behavioural intention. The research model is illustrated in Fig. 5.

The key questions asked in the interview were related to their general feelings of the experience of using the hybrid system, the knowledge they acquired when accessing heritage contents, their expectations of the technology use, and interaction limitations which failed their expectations. Some specific questions were asked to further understand the observed user reactions and conversations during the experience, such as a specific movement or comments they made. Users were also invited to provide suggestions and proposals for future improvements.

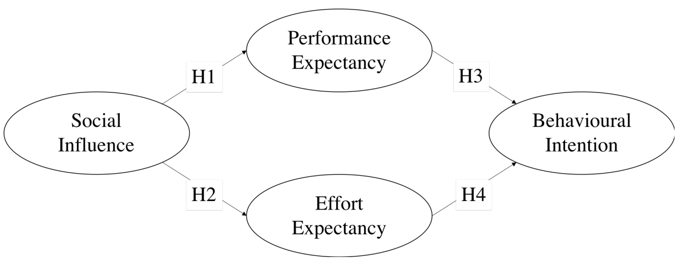


Fig. 5. Our research model, adapted from UTAUT [5]

2) Data Analysis

The data collected from questionnaires were analysed and visualised using SmartPLS to test the reliability of each construct, and to identify the correlations and regression patterns between constructs for statistical significance. Qualitative data from observation notes and interviews used thematic analysis, from which narratives and the thoughts of users were understood. The NVivo software were used to efficiently code the observation and interview notes, subsequently examining repetitions, conflicting answers, relationships, and to generate themes through 'identifying and describing both implicit and explicit ideas within the data' (see [23], [24] for the established method). Analysed results and meanings extracted from the data were compared with our initial hypotheses.

IV. Results

A. Pre-experiment Questionnaire

Our pre-experiment questionnaires asked questions about users' prior experience with 3D gaming, VR, and AR. We found that 63% of our users had 3D gaming experience; 62% of the users had VR experience; and 42% of the users had AR experience. Users experienced with 3D gaming are skilful at them. Users are less skilful with VR and AR overall (see Fig. 6). Our findings demonstrated that length of time users had with 3D games is positively correlated with their perceived 3D gaming skills, which also relates positively with their skills with VR and AR (see Table III).

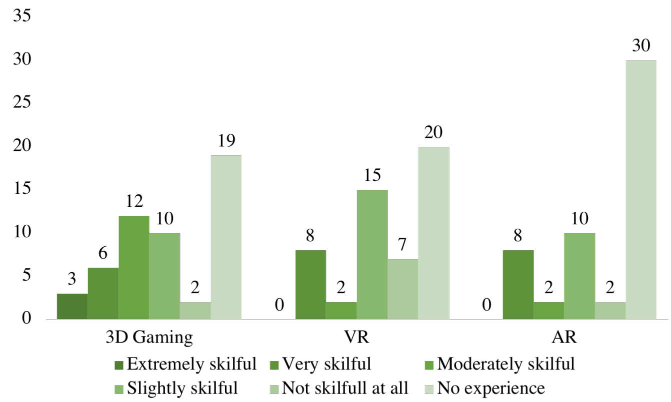


Fig. 6. Self-reported (perceived) skills in 3D gaming, VR, and AR

TABLE III. Correlations Between 3D Gameing, VR, and AR

	3D Gaming Time	**3D Gaming Skill**	**VR Skill**	**AR Skill**
3D Gaming Time	1			
3D Gaming Skill	0.85	1		
VR Skill	0.40	0.55	1	
AR Skill	0.56	0.57	0.27	1

Although most 360 videos are not interactive and sometimes not considered as VR, they can provide an immersive experiences. We therefore listed it as an option for users for claiming to have prior VR experience. Half of the VR users (50.0%) reported to have watched 360 'VR' videos or movies; In our data, most VR users have played VR games (59.4%), while few have experienced VR-based exhibits or educational apps (28.1%). The devices they used were mostly HTC Vive (59.4%) and Samsung Gear VR (53.1%), followed by the Oculus Rift (21.9%) and Google Cardboard (6.3%). Most AR users used it for games (81.8%) and maps (72.7%). Some used AR for marketing campaigns (40.9%) and as tour guide systems (18.2%). In addition, 13% of all users responded that they have had VR or AR experience with other users. The examples they provided were playing Pokémon Go and watching VR movies together with their friends. None has had experience with hybrid VR and AR systems.

Knowledge of museums and cultural heritage were shown to be average in our users. No one considered them extremely knowledgeable or very knowledgeable. 46.2% of users perceived that they were moderately knowledgeable, 48.1% of users rated slightly knowledgeable and 5.8% of users rated themselves to be not knowledgeable at all. However, interest is there. When users see interesting objects in museums, they tend to read the information label (73.1%), take a picture (73.1%), and engage in some social activities, such as asking someone to join in for discussions (34.6%). Users also send private message to friends or posts on social media (40.4%). 34.6% of our participants reported that they would search for more information on the Internet. This implies that users are naturally inclined to acquire and share information. To our participants, the museum visiting experience was more than a learning process, it was also for social activities both inside and outside the boundaries of museums, which are perceived to be important aspects of their social life.

B. Acceptance Questionnaire

The acceptance questionnaire used in this study was adapted from the UTAUT model [5], using a 5-point Likert-type scale ranging from Strongly Disagree (1) to Strongly Agree (5). The items used are listed in Table V.

In order to validate the instruments and test the hypothesised relationships between social influence, performance expectancy, effort expectancy, and behavioural intention, we performed structural equation modelling (SEM) using SmartPLS [25]. SEM involves the assessment of both measurement model and structural model. The measurement model defines latent variables using multiple observed variables and the structural model describes the relationships between latent variables [26]. In this study, the latent variables are constructs in the research model, namely the social influence, performance expectancy, effort expectancy, and behavioural intention.

1) Measurement Model Assessment

Partial least squares path modelling (PLS-PM) allows the estimation of relationship models with multiple latent variables. The PLS algorithm is a statistical regression method, which is essentially an extension of the multiple linear regression model, evaluating the relationship between several independent variables and a dependent variable. It is particularly useful when there are more predictors than observations. The PLS-PM was applied to generate the estimation of item loadings and path coefficients [27], [28]. The results are shown in Table IV.

The Cronbach's Alpha, Composite Reliabilities (CR) and Average Variance Extracted (AVE) are indicators of the construct reliability and validity. Cronbach's Alpha measures the internal consistency, that is, how closely related a set of items are as a group. Items used in different constructs were shown to be robust and reliable measures regarding the Cronbach's Alpha was greater than 0.70 in all cases.

The correlations between the original items and the constructs are illustrated by factor loadings. All factor results of instruments applied in this study were greater than 0.72, which exceeded the suggested factor loading value of 0.50 [26] and met a stricter threshold value of 0.70 [29]. With the factor loadings, the composite reliability and the average variance extracted can be calculated using formula (1) and (2), where λ is the factor loadings and δ is the measurement error, $\delta_i = 1 - \lambda_i^2$.

$$\mathrm{CR} = \frac{\left(\sum_{i=1}^{n}\lambda_i\right)^2}{\left(\sum_{i=1}^{n}\lambda_i\right)^2 + \left(\sum_{i=1}^{n}\delta_i\right)} \tag{1}$$

$$\mathrm{AVE} = \frac{\sum_{i=1}^{n}\lambda_i^2}{n} \tag{2}$$

The composite reliabilities of the constructs ranged from 0.85 to 0.91 and the average variance extracted ranged from 0.58 to 0.75, which exceed the recommended threshold values of 0.70 and 0.50 respectively [26]. Therefore, the constructs in the model have sufficient reliability and convergent validity.

2) Structural Model Assessment

Since PLS-PM does not assume a normally distributed data, which disqualifies the parametric tests for significance. Therefore, the bootstrapping procedure was applied to perform non-parametric tests to obtain the statistical significance of the path model coefficient [30]. The results showed that the model fits the study data as indicated by estimates, R-squared, and significance values (see Fig. 7 and Table VI.).

Estimates show the impact of a relationship between two variables, and the p-value indicates how significant the relationship is. All paths in the model are shown to be statistically significant ($p < 0.05$) and the model fits the overall data ($R^2 = 0.55$).

All four hypothesis were found to be supported. H1 was supported, indicating that the social influence has a positive effect on performance expectancy. H2 was supported, indicating that the social influence has a positive effect on effort expectancy. H3 was supported, indicating that the performance expectancy has a positive effect on behavioural intention. H4 was supported, indicating that the effort expectancy has a positive effect on behavioural intention.

TABLE IV. RELIABILITY AND VALIDITY OF RESEARCH VARIABLES

	Indicator	Factor Loadings	Cronbach's Alpha	CR	AVE
PE	PE1	0.87	0.84	0.89	0.68
	PE2	0.82			
	PE3	0.84			
	PE4	0.76			
EE	EE1	0.82	0.86	0.91	0.71
	EE2	0.89			
	EE3	0.85			
	EE4	0.81			
SI	SI1	0.72	0.76	0.85	0.58
	SI2	0.78			
	SI3	0.76			
	SI4	0.78			
BI	BI1	0.87	0.83	0.90	0.75
	BI2	0.89			
	BI3	0.84			

TABLE V. ADAPTED UTAUT QUESTIONNAIRE ITEMS

Constructs	Question Description	Mean	Std Deviation	N
Performance Expectancy (PE)	PE1: The hybrid VR and AR system is very useful in presenting cultural heritage objects.	4.29	0.85	52
	PE2: The hybrid VR and AR system is very useful in enabling me to get access to more artefacts.	4.27	0.84	52
	PE3: The hybrid VR and AR system can increase my chances for getting more knowledge.	4.29	0.70	52
	PE4: The hybrid VR and AR system can increase my productivity.	3.87	1.10	52
Effort Expectancy (EE)	EE1: Learning how to use the hybrid VR and AR system is easy for me.	4.46	0.80	52
	EE2: Interaction with the hybrid VR and AR system is clear and easy to understand.	4.44	0.73	52
	EE3: It is easy for me to become skilful at using the hybrid VR and AR system.	4.29	0.82	52
	EE4: I like the visiting experience with the hybrid VR and AR system.	4.35	0.88	52
Social Influence (SI)	SI1: People who influenced my behaviour, such as my friends think I should use this system.	3.81	0.84	52
	SI2: People who are important to me think that I should use this system.	3.69	0.90	52
	SI3: I think that the museums and galleries will adopt hybrid VR and AR system in the future.	4.29	0.80	52
	SI4: In general, I can get enough support for the usage of hybrid VR and AR system.	4.06	0.92	52
Behavioural Intention (BI)	BI1: I intend to use the hybrid VR and AR system within the next 12 months if accessible.	4.04	0.88	52
	BI2: I assume that I will be using the hybrid VR and AR system within the next 12 months if accessible.	4.00	0.79	52
	BI3: I plan to use the hybrid VR and AR system within the next 12 months if accessible.	3.98	0.80	52

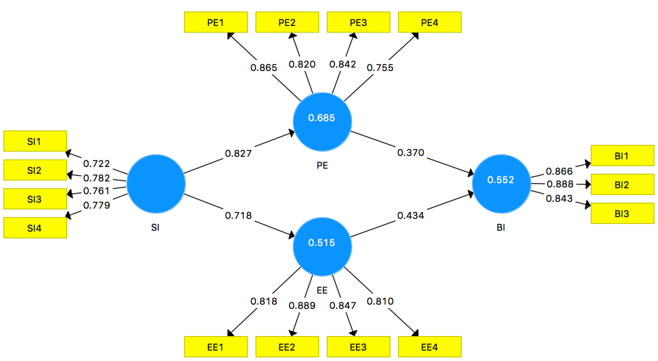


Fig. 7. Research model results

TABLE VI. RESULTS OF HYPOTHESIS TESTING

Hypothesis	Estimates	*t*-value	*p*-value	Confirmed
H1: SI -> PE	0.83	12.32	0.00	Yes
H2: SI -> EE	0.72	7.43	0.00	Yes
H3: PE -> BI	0.37	2.60	0.01	Yes
H4: EE -> BI	0.43	3.53	0.00	Yes

C. Observations and Interview

General feedback about the hybrid VR and AR experience was positive. VR users responded that the environment was immersive and the objects looked real. The object-grab interaction was intuitive and helpful as it allowed them to observe objects from different angles, which is not possible in museums considering the risk of breaking the artefacts. Some also compared this interactive VR experience with their previous 360 video experience and responded that the interactions engaged them better and were fun to play with. AR users acknowledged the great quality of the models and reported that seeing the 3D models of artefacts augmented on top of the 2D images on their cube surfaces felt like the artefacts came alive. They also reported that holding the cube with digital artefacts is comparable to holding the objects in hand. The interaction with 2D images on the sides of the cube helped support the user mental model that the smartphone camera can transform the 2D image of a cultural artefact into an equivalent 3D model. The use of a physical cube of appropriate size provided users with physical embodiments of the artefacts and gave users the impression that the cube embeds the artefacts and that the smartphone camera is the key to unlock them.

During the experiment, most interactions between the VR user and the AR user occurred around our virtual objects, which achieved our goal of object as interface. In most cases, VR users noticed the changes in the virtual environment, and asked "Are you looking at the Tri-coloured Camel (object #5, see Table I.)?" when they saw an object rotating. Moreover, most of them walked towards the object and looked at it with their paired partner if they noticed the change. We have also found AR users prompting the VR user to engage with an object of interest together. Although the virtual environment was a simple room without any historical context, and AR users only had access to 3D models and information label of the objects, the objects they shared became the interface between them, connecting them and mediating their communications. We found that most discussions were initiated by users' observation of the appearance of the artefacts, such as their shape and colours. Users were seen to develop conversations through the sharing of their interpretations of the objects as well as their associated knowledge from literatures or from visiting museums. A user who first saw the 3D Bronze Mask (object #2, see Table I.) initiated a discussion, explaining to his partner about the history behind the artefact because it reminded him of the *Sanxingdui* documentary film he had previously watched. Some related the artefacts with stories they have heard from their friends, as well as news they have read elsewhere. For example, a girl narrated a piece of news she read and said how she wished the Pottery Unicorn (object #4, see Table I.), as a symbol of justice, could help punish the bad guy in the news. With the shared objects provided as a communication interface, users had generated interesting contents during the experience, both within and beyond the scope of the cultural heritage context we have provided in our experiment. This suggested that despite the different technology used and the limited information provided, the 3D objects themselves were mediating communications.

Although the control methods were identical across our 26 pairs of users, some were creative with their interactions. Because the Figure of an Assistant to the Judge of Hell (object #6, see Table I.) was a statue of 5 feet, we deliberately made it immovable in the virtual environment, which made it difficult for the VR user to observe its back. However, some VR users asked their AR partner to observe it, triggering a rotation so that they could see it from the back. This spontaneous cooperative interaction illustrated how users can become creative in taking advantage of the asymmetrical access to engage with the system and the other users. It also provided inspirations for our future design, incorporating asymmetrical access to foster user interaction, engagement, and communication.

In another example, VR users switched the positions of artefacts in the virtual environment, and put them back to the original positions before the experiment was completed. When being asked about the action, they reported that they initially explored the possibility of placing the object randomly within the environment but failed. However, they realised that the objects could switch places, i.e., pedestals, and being courteous, they placed the objects back at their original locations so that new users will not be confused. This demonstrated that users were willing to explore the affordances of the environment and that they could adapt to it quickly. More importantly, users took the virtual environment as a social environment, and that they were expecting others to see what they saw. Such observed behaviour should be supported and made used of in the future design of multiuser virtual environments if we intend to include a community for supporting, moderating and maintaining user-generated content.

Social presence was also reported in our experiment. We observed that there were infrequent interactions between a pair of users who were strangers to each other, they reported that although they did not have the intention of approaching the other user, they had the sense that, through the cues given by the object, they were accompanied by another person during their experience. The awareness of another user through the object as an interface mitigated loneliness for them in a fully immersive environment.

D. Limitations

Our current design of the hybrid VR and AR system is not without limitations as it is preliminary and exploratory. Therefore, limitations and underestimation of the experience is expected. Here we report on the limitations of the environment, which will inform our future design of our pioneering system.

We observed that there were interactions which failed user expectations. Some VR users attempted to affect the AR user by making the artefact disappear or move when they picked up the object on their side, which is the virtual environment. During the interview, our users reported that it would have been more interesting if they could affect their partner's world via some interaction mechanisms. One of the advantages which users highlighted, for the reasons they liked VR and AR technologies were that these technologies can allow more dynamic interactions as compared to physical museum. The use of such devices is expected for future applications for cultural heritage. Users have also suggested that games should be introduced to promote learning, such as categorising artefacts based on materials or sorting them according to historical, chronological sequence. A user narrated that when he found that he could not move the Figure of an Assistant to the Judge of Hell (object #6, see Table I.), his immediate thought was to work together with his partner to try to lift it up. These reflections supported the idea that the future design of hybrid VR and AR system should include more flexible affordances. Research has suggested that game design elements should be integrated to provide a playful and game-oriented environment, which corroborates the idea of gamification, empowering motivation, engagement, and loyalty [31]. We have also discovered that most users came to the conclusion that the interactions with objects and with their partners contributed to their understanding and interpretation of cultural heritage. The social interaction within the environment encouraged paired users to observe more and read more, culminating in more knowledge gained of the contents within the environment and strengthening their memory.

V. Conclusion

This research explored the use of a hybrid VR and AR system in a multiuser cultural heritage application, and confirmed the proposed research model on acceptance. We demonstrated, through our analysis of our datasets from 52 participants in 26 pairs, that users' behavioural intention depends on performance expectancy and effort expectancy, which are positively affected by social influence. Users considered the technology we have developed helpful in giving access to cultural heritage objects. They have found that the interface was easy to use. The understanding of the social nature of users and developing technology in supporting it is fundamental to manage their expectations resulting in the achieving of higher behavioural intentions. Our study has confirmed that social influence is important, and should be taken as a significant factor in the future design of multiuser systems supporting different devices and interface features.

In addition, the social nature that users possess in their actual museum visiting experience can be transferred to the virtual environment – their interaction and communication were mediated by cultural artefacts. Many users took museum visiting as a social activity involving families and friends. Users also engaged in social activities outside the boundary of the museum, such as the sharing of their experience via social media. Our observations and interviews confirmed such a social nature and supported the idea that the virtual heritage objects could be an interface between their interaction and communication. Interactions around digital interfaces have been shown to be useful to promote new topic of discussion and foster communication [32]. Our study further showed that digital objects, together with the discussions and communication that users initiated themselves assisted with the interpretations of the artefacts, facilitated by the virtual environment and the interactions afforded by the different devices. Furthermore, users perceived social presence within the virtual environment and felt a sense of immersion and accompaniment. They acknowledged the positive influence of their interactions with others on their experience and expected more interaction possibilities and closer connections with other users. Users have expressed that such a system should include game elements in the future design of hybrid VR and AR systems. Our findings also indicated a positive relationship between 3D gaming experience and skills in VR and AR, although the experiment showed that the control methods were intuitive enough for first time users to learn and master quickly. This may have explained their expectations for richer interactions and the presence of game elements.

The findings contributed to both the theoretical and methodological frameworks which will help guide the future design of multiuser digital systems for cultural heritage and 'museums outside the walls' facilitating the social nature of museum visits. However, the results should be interpreted carefully due to these observed limitations. First, the majority of users in our study participated with people they knew beforehand. The actual interaction, engagement, and communication between strangers will need to be studied. Second, the majority of our users were university students, aged between 18 and 35. Therefore, the results may not be applicable for a general museum visitors. However, this indicates an opportunity to attract younger visitors. We felt that a more comprehensive study is required to understand users of other age groups, such as family units with children. However, the younger users in this study are more likely to be open to new technologies and to have their own VR and AR devices in the future. The understanding of their acceptance and behaviour is nevertheless significant for future VR and AR system design in cultural heritage.

Acknowledgement

This work was carried out at the NVIDIA Joint-Lab on Mixed Reality. The authors acknowledge the financial support from IAMET, Ningbo Education Bureau, Ningbo Science and Technology Bureau, Zhejiang Provincial commonwealth grant, China's MOST and the University of Nottingham. This work is part of the International Doctoral Innovation Centre (IDIC) programme and partially supported by EPSRC grant no EP/L015463/1.

References

[1] J. Forlizzi and K. Battarbee, "Understanding experience in interactive systems," *Conf. Des. Interact. Syst.*, pp. 261–270, 2004.

[2] E. Ch'ng, "Experiential Archaeology: Is Virtual Time Travel Possible?," *J. Cult. Herit.*, vol. 20, no. 2009, pp. 458–470, 2009.

[3] E. Ch'ng, "The Mirror Between Two Worlds: 3D Surface Computing Interaction for Digital Objects and Environments," in *Digital Media and Technologies for Virtual Artistic Spaces*, Hershey, PA: IGI Global, 2013.

[4] F.-T. Leow, E. Ch'ng, T. Zhang, S. Cai, and S. See, "'In-The-Wild' Observation and Evaluation of a Chinese Heritage VR Environment with the HTC VIVE," in *International Conference on Virtual Systems and Multimedia (VSMM), 31 Oct – 2 Nov*, 2017.

[5] V. Venkatesh, M. Hall, G. B. . Davis, F. D. . Davis, S. M. Walton, and M. G. . Morris, "User Acceptance of Information Technology: Toward a Unified," *MIS Q.*, vol. 27, no. 3, pp. 425–478, 2003.

[6] I. E. Sutherland, "The Ultimate Display," *Invit. Lect. IFIP Congr.*, vol. 65, pp. 506–508, 1965.

[7] E. Ch'ng, "Experiential Archaeology: Is Virtual Time Travel Possible?," *J. Cult. Herit.*, vol. 20, no. 2009, pp. 458–470, 2009.

[8] National History Museum, "Explore virtual reality with David Attenborough's First Life," 2015. [Online]. Available: http://www.nhm.ac.uk/discover/news/2015/june/dive-back-in-time-with-david-attenborough-s-first-life.html. [Accessed: 06-Mar-2018].

[9] E. Ch'ng, R. J. Stone, and T. N. Arvanitis, "The Shotton River and Mesolithic dwellings: recreating the past from geo-seismic data sources," in *The 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage, VAST04: Interdisciplinarity or "The Best of Both Worlds": The Grand Challenge for Cultural Heritage Informatics in the 21st Century*, 2004, pp. 125–134.

[10] T. Lambert, *Stone Age Atlantis, National Geographic Television*. National Geographic Television, 2010.

[11] V. Gaffney, S. Fitch, E. Ramsey, R. Yorston, E. Ch'Ng, E. Baldwin, R. Bates, C. Gaffney, C. Ruggles, T. Sparrow, A. McMillan, D. Cowley, S. Fraser, C. Murray, H. Murray, E. Hopla, and A. Howard, "Time and a place: A luni-solar 'time-reckoner' from 8th

millennium BC Scotland," *Internet Archaeol.*, vol. 34, no. 34, 2013.

[12] M. Roussou, L. Pujol, K. Akrivi, i C. Angelik, P. Sara, and M. Vayanou, "The museum as digital storyteller: Collaborative participatory creation of interactive digital experiences," *MW2015 Museums Web*, p. 210, 2015.

[13] E. Ch'ng, "New ways of accessing information spaces using 3D multitouch tables," in *Proceedings of the 2012 International Conference on Cyberworlds, Cyberworlds 2012*, 2012, pp. 144–150.

[14] A. Katifori, S. Perry, M. Vayanou, L. Pujol, A. Chrysanthi, and Y. Ioannidis, "Cultivating mobile-mediated social interaction in the museum: Towards group-based digital storytelling experiences," *MW2016 Museums Web 2016*, pp. 1–14, 2016.

[15] F. D. Davis, "A Technology Acceptance Model for Empirically Testing New End-User Information Systems," 1985.

[16] A.-C. Haugstvedt and J. Krogstie, "Mobile augmented reality for cultural heritage: A technology acceptance study," in *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2012, pp. 247–255.

[17] M. Carrozzino and M. Bergamasco, "Beyond virtual museums: Experiencing immersive virtual reality in real museums," *J. Cult. Herit.*, vol. 11, no. 4, pp. 452–458, 2010.

[18] P. Zhang, "Technical opinion: Motivational affordances: reasons for ICT design and use," *Commun. ACM*, 2008.

[19] J. M. Holden and J. B. Overmier, "Choice behavior under differential outcomes: Sample stimulus control versus expectancy control," *Learn. Motiv.*, 2015.

[20] Vuforia, "SDK Download | Vuforia Developer Portal," *developer.vuforia.com*, 2018. .

[21] J. Gugenheimer, E. Stemasov, J. Frommel, and E. Rukzio, "ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, 2017, pp. 4021–4033.

[22] P. Dourish, "Where the Action Is: The Foundations of Embodied Interaction," *Where the action is the foundations of embodied interaction*, vol. 36, no. 3. p. chaper 1-3, 2001.

[23] G. S. Guest, E. E. Namey, and M. L. Mitchell, *Collecting qualitative data; a field manual for applied research*. 2013.

[24] W. J. Gibson and A. (Andrew J. Brown, *Working with qualitative data*. SAGE, 2009.

[25] C. M. Ringle, S. Wende, and A. Will, "SmartPLS 3.0," *Http://Www.Smartpls.De*, 2005.

[26] J. F. Hair, W. C. Black, B. J. Babin, R. E. Anderson, and R. L. Tatham, "Multivariate Data Analysis," *Prentice Hall*, 2010.

[27] J. F. J. Hair, G. T. M. Hult, C. Ringle, and M. Sarstedt, *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. 2014.

[28] J. Henseler, C. M. Ringle, and M. Sarstedt, "Using partial least squares path modeling in advertising research: basic concepts and recent issues," *Handb. Res. Int. Advert.*, 2012.

[29] C. Fornell and D. Larcker, "A second generation of multivariate analysis: Classification of methods and implications for marketing research," *Review of marketing*, vol. 87. pp. 407–450, 1987.

[30] A. C. Davison and D. V Hinkley, "Bootstrap Methods and their Application," *Technometrics*, 1997.

[31] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: Defining gamification," *Proc. 15th Int. Acad. MindTrek Conf. Envisioning Futur. Media Environ. - MindTrek '11*, 2011.

[32] C. Rocchi, D. Tomasini, O. Stock, and M. Zancanaro, "Fostering conversation after the museum visit," *Proc. Work. Conf. Adv. Vis. interfaces - AVI '08*, p. 335, 2008.