On the Relationship between Subjective and Objective Measures of Virtual Reality Experiences: a Case Study of a Serious Game

Inmaculada Rodríguez, Anna Puig
Dep. Matemàtiques i Informàtica, University of Barcelona, UBICS, IMUB
Barcelona, Spain
{inmarodriguez,annapuig}@ub.edu

Juan Antonio Rodríguez-Aguilar, Josep Lluís Arcos, Sergi Cebrián Anton Bogdanovych
IIIA-CSIC, Campus UAB, University of Western Sydney
Cerdanyola, Spain
{jar,arcos,scebrian}@iiia.csic.es, A.Bogdanovych@westernsydney.edu.au

Núria Morera, Raquel Piqué Antoni Palomo
UAB, Museu d'Arqueologia de Catalunya
Barcelona, Spain
{nuria.morera,raquel.pique}@uab.cat, antoni.palomo@gencat.cat

ABSTRACT
In this paper we present a Virtual Reality game related to Cultural Heritage. We contribute with an analysis of subjective measures taken from questionnaires filled by users after the VR experience, and objective measures taken from logs during the VR game. Specifically, we were interested on study data globally and in groups of user behaviour. Analysing data globally we see a high value of users' subjective perceptions. Nevertheless, we found differences of subjective measures when splitting the Novice group.

Specifically, the subjective perception of Strugglers is considerably lower than the rest of groups, and this difference is significant. Then, we propose strategies to provide a better experience to Strugglers. We also found correlations between objective and subjective data when they were analysed globally (i.e. without using groups), but these measures did not correlate when they were analysed using behaviour groups.

Author Keywords
serious game, cultural heritage, virtual reality, VR experience

INTRODUCTION
Nowadays Virtual Reality (VR) experiences are designed to be more and more interactive and engaging. Thus, its widespread use along different areas such as education, health, entertainment and cultural heritage. Concretely, VR initiatives in the cultural context help curators to transfer historical knowledge in a different but complementary way to classical expositions. Moreover, the interactivity of VR applications and their novelty attract not only young visitors but also the general public.

In: J. Arnedo-Moreno, C.S. González, A. Mora (eds.): Proceedings of the 3rd International Symposium on Gamification and Games for Learning (GamiLearn'19), Barcelona, Spain, 22-10-2019, published at http://ceur-ws.org

To date, cultural experiences in VR have been designed either as walks around virtual reconstructions where the visitor immerses as a mere spectator [1] [12] or as interactive experiences where the visitor plays an active role [20] [19]. We are interested in the second type of experiences with an additional ingredient, which is the challenge of a game. The goal of this game is to provide visitors with both fun and learning, namely Informal Learning (IL). IL is an instructional activity that is not usually a part of the standard curriculum of educational institutions. It provides the users with knowledge and understanding of a variety of subjects in museums, science centres, exhibitions, even in the workplace [2]. However, to know the actual learning achievement is not an easy task not just due to the informal context but also due to the difficulty of managing VR interactions.

This research is situated in the context of Cultural Heritage dissemination, and presents a VR game about La Draga, a Neolithic settlement located in Catalunya, Spain. In a previous work we presented results of an Informal Blended Learning experience (workshop + exhibition + a 360°video + a VR game) aimed at the general public (families, usual visitors, tourists) [18]. There we collected quantitative measures of VR interactions through visitors' logs. Then, we performed data analytics for exploring patterns of visitors' behaviours.

In this paper we focus on one piece of the aforementioned Blended Informal Learning experience, the VR game, and contribute with a comprehensive analysis of subjective measures taken from questionnaires filled by users after the experience, and objective measures taken from logs during the game. We analyse users' subjective measures assessing their interactive and learning experience in VR, first globally, and then grouped in the aforementioned user behaviour patterns. We also explore the relationship between users' subjective measures with their corresponding objective ones. Specifically, the correlation between perception of easiness of movement and number of teleports. Finally, we examine users' learning perception on the defined behaviour groups.

Our results show different scenarios when we analyse the sample without and with groups. Meanwhile we could considered users' subjective scores particularly positive when they are analysed globally (i.e. without using groups), they manifest differently when the analysis is done in groups. Consequently, we propose design strategies in the direction of adapting VR experience content to improve perceptions of those groups of users that showed room for improvement.

RELATED WORK

In the literature we find different serious game initiatives for Cultural Heritage (CH). Attending to learning objectives of serious games, Mortara [17] proposed a taxonomy considering three categories: cultural awareness, historical reconstruction, and heritage awareness. First, cultural awareness, which is concerned with intangible cultural heritage such as customs, folklore, and rules of behaviour in an ancient society. Several games situate in this category. For instance, games Papakwaqa [14] and Icura [9] tackle cultural issues (e.g beliefs, ceremonies) in Taiwan and Japan respectively. Second, the category of games named historical reconstruction teaches history through reconstructions of historical periods and events. These games are usually designed as strategy [22] and role-play games, where the player has an active role in the game, e.g as the soldier who fights in an historical battle. Examples of games in this category are The Battle of Thermopylae, The Siege of Syracuse, and The ancient Olympia [7] [6] [10]. And third, heritage awareness category comprises serious games where users immerse in a virtual reconstruction of historical buildings and locations, with the aim of learning about them. Specifically, in the area of archaeology, "Multi-touch Rocks" [21] is a game about engravings in caves in France and Italy. Players used a multi-touch tabletop to explore a big image of a rock, from there they accessed to mini-games about hunting oxes, building a house, etc. Our proposed game, La Draga, also belongs to this category. We recreate the Neolithic site so that visitors could feel and understand the settlement, hopefully arousing their curiosity to go to the physical site.

Up to now we have explored serious games in a general context of Cultural Heritage. Next, we focus on CH solutions using VR technology. Several research works considered VR technology as supporting learning and social behaviour in museums [8] [4]. Related to interaction and immersion in VR, Carrozino [3] proposed a classification of VR installations for cultural heritage. On the interaction dimension, the author classified systems from non-interactive, to using mediated interaction (device based interaction), and up to using natural interaction (gestural interaction, speech recognition). Regarding immersion, systems were arranged in three groups: non-immersive, low-immersion, and high-immersion, and were analysed based on visual, acoustic, haptic and motion features. We situate La Draga game in the category of high immersive systems (room-scale Head Mounted Displays) but with medium level of interaction (dependent on devices that the visitor wears, i.e the HTC Vive controllers). Other research works that can be also considered high immersive are the virtual walk of the Greek city of Miletus [11], and the 3D vignettes of Viking settlements (scenes, objects and activities) in Ireland [20]. Nevertheless, these works did a low level use of interaction, as they mainly focused on 360°videos or 3D scenes with fixed viewpoints.

Usually, VR experiences have been evaluated through post-questionnaires. A research work [5] explored the possible replacement of these post-questionnaires by task performance measurements, although they concluded that subjective post-tests can still be more reliable than evaluating task performance. Following this idea, we propose to explore the relationship between subjective and objective measurements taken into account different user's behaviours.

LA DRAGA GAME

Figure 1. The Virtual Draga environment.

La Draga game takes place in the Early Neolithic era (5.300-4.800 BC) on La Draga settlement. This settlement is located at the shores of the Banyoles Lake at the NE of the Iberian Peninsula. Its relevance lies on the high degree of preservation of archaeological remains such as vegetal fibres and wooden tools. In the museum's physical exhibition space, visitors learned through watching the physical remains and some reconstructed hunts. La Draga game puts these remains in context, which is a trustworthy virtual environment reconstructed according to the archaeologists' criteria (see Figure 1), enabling users not only to explore the outside environment and hunts, but also to be able to find and interact with objects previously seen in the Museum.

Thus, the virtual Draga landscape includes huts, inhabitants, animals, vegetation, etc. in an open world landscape, as well as objects and remains in indoor hunts (see Figure 2).

Figure 2. Outdoor and indoor details of La Draga game.

Game players act as time travellers, who visit the Draga settlement. Their mission is to walk around the environment and find Non-Neolithic objects, and throw them to a big object, from another era, that represent the time portal, which devours all Non-Neolithic objects and repels Neolithic ones.

Figure 3. Landing area, and time portal as a blue sphere. Initially, near to the landing area, player can find quickly a highlighted "Cola can".

Game mechanics are simple. The player lands near to the temporal portal (Figure 3). Just to facilitate the *on-boarding*, a Non-Neolithic object is located nearby. The player walks around the environment. When she sees a Non-Neolithic object (i.e a Cola can), grabs it, and throws it into the portal. If done well, that is, the object enters the portal and the object belongs to the Non-Neolithic era, the player scores 100 points. If done incorrectly, i.e it is a Neolithic object, she loses 50 points. The game lasts five minutes. This *time constraint* is highly related to the deployment of games in museums [13], i.e. to avoid because too many people queuing and waiting for playing.

We also include some quests and surprises along the game. For instance, players should distract a wild pig to pick up a "Cola Can" (a Non-Neolithic object that player should throw into the portal), by throwing apples, and thus shifting the pig's focus away from the "Cola Can". Here, players should be fast enough picking up the can so not to get caught by the pig (see left picture in Figure 4). Moreover, players could activate some *Easter eggs*, which are Neolithic inhabitants' animations. For instance, picking up a Neolithic sickle located near a hunt, players can see Neolithic inhabitants entailed in woodworking (see on the right in Figure 4).

Figure 4. Surprises along the game: the pig's challenge and the virtual human animation in the Draga Game.

Actually, while exhibition visitors were in line they could saw another visitor playing the game in a monitor TV (see Figure 5). The use of the monitor is an usual practice in VR exhibitions as it allows visitors to get a *previous idea of the VR experience*. We did not want visitors waiting in line to have advantages because they could saw where objects were located. For that reason, 5 Neolithic and 6 Non-Neolithic objects were randomly placed in the game level. The Neolithic objects were: a Bow, a Sickle, a Ladle, an Adze and Apples as pig's food. The Non-Neolithic objects were: a Saw, A Cola Can, Bananas, a modern Bow, a Knife, and Darts.

Players used the VR controllers to interact with La Draga world. The trackpad allowed them to *move* by locating the target of the movement. This target is depicted as the end point of a parabola (see Figure 5). Also, users *grab* objects pulling the trigger control on selectable objects. Finally, *throwing objects* implies to release the trigger control meanwhile doing the gesture of throwing. As they are non trivial interactions, we included a *training tutorial* to help players to learn early the usage of the VR controllers. Moreover, we added different types of feedback: visual (i.e. highlighting selectable objects), audio (i.e. when an object enters to the time portal) and haptic feedback (i.e. when picks attacked, player can feel pressure on the controls).

We employed Blender[1] to model the 3D environment, objects and animations, under the supervision of archaeologists. We developed La Draga game using Unity[2] with a VR setup. Finally, users play the game using an HTC VIVE headset with wired HMD (Head Mounted Display) and wireless hand-controllers (see Figure 5).

Figure 5. VR Settings: Teleport.

EVALUATION

Participants and method

We recruited 49 voluntary users that attended the EAA (European Association Archaeology) conference. Users main profile was academics, practitioners and researchers on the archaeology field. The subjects were 55% male and 45% female.

A post-experience questionnaire was filled by the users. It consisted of questions related to demographics and their expertise in VR and playing games. Users expertise in games and VR is depicted in Figure 6[3]. It is asymmetrically distributed, with a high number of users having no previous experience neither in playing games nor VR. Data also show that experienced users

[1] http://www.blender.org

[2] http://docs.unity3d.com

[3] Data were transformed from the range between 0 and 10 to four categories (Almost nothing, Somewhat, Enough, A lot)

were more in playing games than in VR technology, 45,09% of subjects scored their game expertise as "Enough" and "A lot", 68,62% of then scored their VR expertise as "Almost nothing" and "Somewhat".

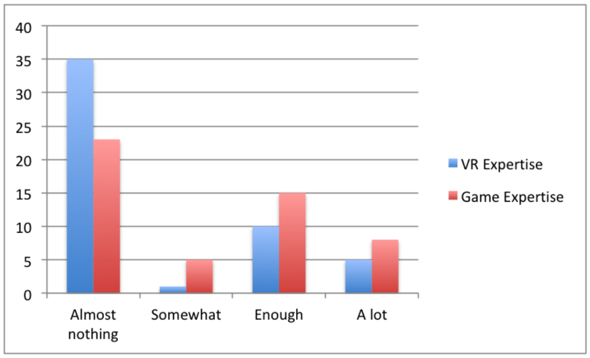

Figure 6. VR and Game Expertise profiles.

Moreover, we included four questions regarding the (subjective) perception of easiness while interacting in VR, and the overall opinion of the experience. All answers ranged from 0 to 10 (see Table 1).

Table 1. Closed-ended questions of the post-experience questionnaire.

Q1	Was it **easy to move around the environment**?
Q2	Was it **easy to throw** objects?
Q3	Was it **easy to find** Neolithic objects?
Q4	**Score the VR Draga** experience

In this study, the objective data comes from our previous work [18], where we identified three patterns of user behaviour (Novices, Medium and Pro), analysing logs from users while playing the VR Draga game. *Novice* players scored very few points and rarely teleported (35.58% of users). *Medium* players achieved medium game score, performed a low-to-medium number of teleports and they were efficient on grabbing objects (41.66% of users). Finally, *Pro* players (18.75% of users) obtained the highest scores, performed a medium number of teleports, and their grabbing efficiency was high. We also found four subpatterns related to the Novice behaviour based on variables like *Number of Visited Zones, Score, Number of Teleports, Grabs, Failed Throws*, and *Pig and Human Interactions*. The four patterns are: *Struggler* (users who really had problems moving and interacting with object), *Engaged* (users that preferred to spend the time with the pig, or watching human animations, and thus discarded to play the game), *Explorer* (users who also discarded to play and preferred to visit the virtual settlement), and *Late-Explorer* (those who had interaction issues initially and then, they decided to explore the environment).

Results

First, we analyse the score achieved by users based on their games expertise. As seen in Figure 7 the participants that obtained higher scores were experienced users. Indeed, a Spearman correlation analysis indicates a relationship between expertise in game and efficiency (ρ = 0,410, p = 0.04).

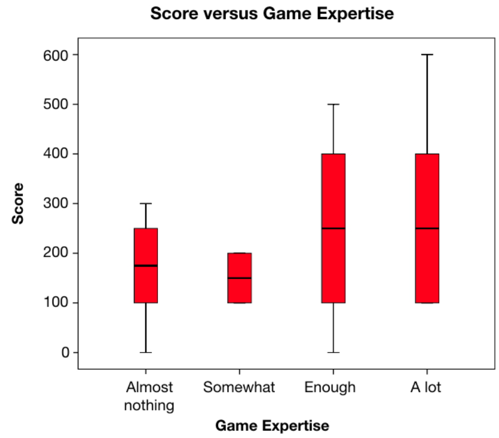

Figure 7. Game score vs Game Expertise.

Regarding users enjoyment and satisfaction with the experience, the question "Score the Draga experience, Q4 in Table 1, obtained an average score of 8.45, which reflects a very positive opinion of users. On the other hand, Figure 8 shows users' perception after the VR experience, i.e. "How easy they found move around the environment, grab and throw objects". Overall, their subjective perception was rather positive. Nevertheless, they found to move in the VR settlement easier than to grab and to throw objects. In fact, the interaction designed to move requires to press just one button of the touch controllers, while grabbing and throwing require to perform a gesture and pressing/releasing the trigger controller synchronously.

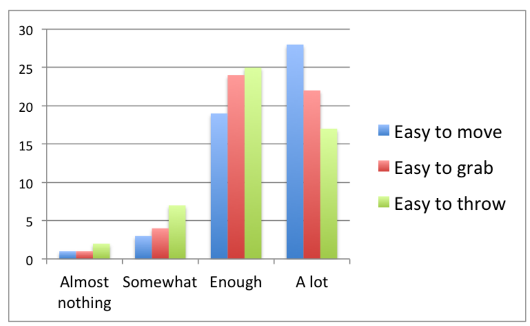

Figure 8. Responses to Q1-Q3 questions.

Specifically, if we focus on analysing objective and subjective measures about movement, we obtain a higher median value on subjective perception (9.0) than on objective measurement (median of normalised number of teleports 3.2584). Actually, these results align with researches that bring into question the reliability of questionnaires and bet for the combination with other analysis and assessment methods to achieve interpretable results [15] [16]. To better analyse these results, we segment the data according to the main behaviour patterns (Novice-Medium and Pro). As shown in Figure 9, users' subjective perceptions about VR interaction (red bars) is rather more positive than the objective data (blue bars), which was measured with the number of teleports.

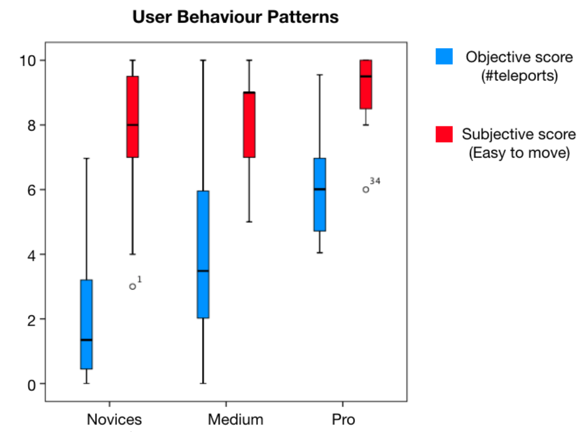

Figure 9. User behaviour pattern vs movement (objective and subjective scores).

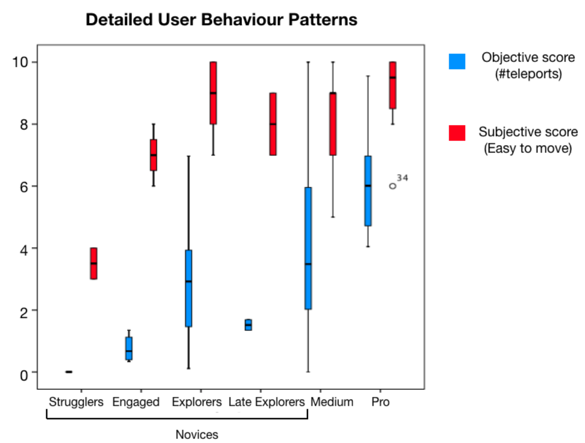

Figure 10. Detailed user behaviour pattern vs movement (objective and subjective scores).

However, we face a different scenario if look in more detail Novices subpatterns (Struggler, Engaged, Explorer, and Late-Explorer) in Figure 10. We can observe that not just the same trend continues between subjective and objective scores, but also other trends emerge. That is, when Novice group splits, the scores corresponding to the four subpatterns change significantly. For instance, the more the users have VR interaction handicaps (Strugglers), the more their perception decreases (see two first bars in Figure 10), showing the lowest values (median of 3.8), not so high as the global Novices scores (median of 8 in Figure 9). Indeed, the Kruskal Wallis no-parametric test (H=10.99;df=5; p-value=0.051) was used to determine if user behaviour patterns influence the users' subjective perception. Using Mann-Whitney to discriminate by pairs, we found the perception of groups Struggler-Explorer, Struggler-Medium and Strugler-Pro differs statistically. Therefore, we could detect Struggler users in run-time, and then help them to better manage movement or VR controllers. Consequently their pattern could evolve to others (Explorer, Medium, Pro), improving their subjective perceptions.

Additionally, we performed a Pearson correlation analysis to explore the relationship between objective and subjective data. Specifically, the number of teleports, as an objective measure, and "Easy to Move perception", as subjective one, correlate when analysing the whole sample (r=0.552, p<=0.001). Again, when we analyse the data based on main behaviour patterns, these measures correlate differently. Actually, Novices (r=0.491, p<=0.001) and Medium players (r=0.598, p<=0.001) follow the same trend than the global sample, meanwhile do not correlate in Pro users. Thus, the initial correlation between these measures in the whole sample could convey us to substitute questionnaires by the number of teleports. However, the no-correlation found in the group of Pro players, prevent us of doing it. Then, it reinforces the necessity of combining subjective data coming from questionnaires and objective data collected through users logs along the experience.

Finally, we asked participants the question "The 3D immersive experience helped you to know how the Draga Neolitic settlement was?" in order to evaluate their learning perception. As seen in Figure11, we found that the highest scores correspond to those patterns that moved easily around the virtual Draga and consequently had a general vision of the landscape, i.e Explorers, Late-Explorers, Medium and Pro users. In fact, note that these patterns did a high number of teleports (see blue bars in Figure 10), more than Strugglers and Engaged users. Again, it reinforces the previous idea of helping Novice users to evolve to other patterns. In this case, an early detection of Novice users on the fly could help to enhance their learning.

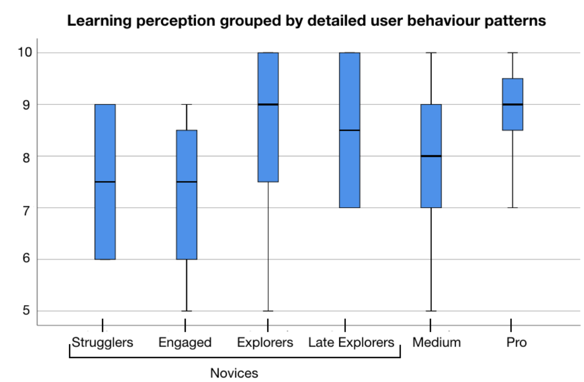

Figure 11. Learning perception by patterns.

CONCLUSIONS AND FUTURE WORK

In this paper we study objective and subjective data gathered during and after a VR experience for cultural heritage dissemination, which is the Draga serious game. When we analysed subjective measures globally, we appreciated high values that may denote they are not reliable, due to several reasons such as the user is impressed with the novelty of the experience, and the user wants to be kind with the evaluator, etc. When we analysed these measures in the six behaviour groups (Strugglers, Engaged, Explorers, Late Explorers, Medium and Pro), we found differences of subjective measures. Specifically, the subjective perception of Strugglers is considerably lower than the rest of groups, and this difference is significant. Thus, detecting Strugglers users during the gameplay could help them to evolve to other patterns, facilitating the VR movement , and in consequence an enhanced perception of the overall experience. On the other hand, we found correlation between subjective and objective measures of motion in VR when analysing globally the data. However, the correlation did not maintain when the analysis was done using the behaviour patterns. As this study has mainly focused on motion measures, and considering we have subjective perceptions about grabbing and throwing, as future work this research will extend the collection and analysis of objective data to these other perceptions.

ACKNOWLEDGMENTS

Projects 2017-SGR-341 and MISMIS-Language (PGC2018-096212-B-C33)

REFERENCES

1. L. Barbieri, F. Bruno, and M. Muzzupappa. 2017. Virtual museum system evaluation through user studies. *J. of Cultural Heritage* 26 (2017), 101–108.
2. Z. Bekerman, N. C Burbules, and D. Silberman-Keller. 2006. *Learning in places: The informal education reader*. Vol. 249. Peter Lang.
3. M. Carrozzino and M. Bergamasco. 2010. Beyond virtual museums: Experiencing immersive virtual reality in real museums. *J. of Cultural Heritage* 11, 4 (2010), 452–458.
4. M. Carrozzino, M. Colombo, F. Tecchia, C. Evangelista, and Massimo Bergamasco. 2018. Comparing Different Storytelling Approaches for Virtual Guides in Digital Immersive Museums. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*. Springer, 292–302.
5. E. Champion, I. Bishop, and B. Dave. 2012. The Palenque project: evaluating interaction in an online virtual archaeology site. *Virtual reality* 16, 2 (2012), 121–139.
6. D. Christopoulos and A. Gaitatzes. 2009. Multimodal interfaces for educational virtual environments. (2009).
7. D. Christopoulos, P. Mavridis, Anthousis A., and J. Karigiannis. 2011. Using Virtual Environments to Tell the Story:" The Battle of Thermopylae". In *2011 Third International Conference on Games and Virtual Worlds for Serious Applications*. IEEE, 84–91.
8. M. Economou and L. Pujol. 2008. Educational tool or expensive toy? Evaluating VR evaluation and its relevance for virtual heritage. In *New heritage: New media and cultural heritage*. Routledge London/NY, 242–260.
9. J. Froschauer, I. Seidel, M. Gärtner, H. Berger, and D. Merkl. 2010. Design and evaluation of a serious game for immersive cultural training. In *2010 16th international conference on virtual systems and multimedia*. IEEE, 253–260.
10. A. Gaitatzes, D. Christopoulos, and G. Papaioannou. 2004. The ancient olympic games: being part of the experience. In *Proc. of the 5th International conference on Virtual Reality, Archaeology and Intelligent Cultural Heritage*. Eurographics Association, 19–28.
11. A. Gaitatzes, D. Christopoulos, and M. Roussou. 2001. Reviving the past: cultural heritage meets virtual reality. In *Proc. of the 2001 conference on Virtual reality, archeology, and cultural heritage*. ACM, 103–110.
12. S. Gonizzi, G. Caruso, LL Micoli, M Covarrubias Rodriguez, G. Guidi, and others. 2015. 3D visualization of cultural heritage artefacts with virtual reality devices. In *25th International CIPA Symposium 2015*, Vol. 40. Copernicus Gesellschaft mbH, 165–172.
13. HT Hou, SY. Wu, PC Lin, YT Sung, JW Lin, and Kuo-En Chang. 2014. A blended mobile learning environment for museum learning. *Journal of Educational Technology & Society* 17, 2 (2014), 207–218.
14. CH Huang and YT Huang. 2013. An annales school-based serious game creation framework for taiwanese indigenous cultural heritage. *J. on Computing and Cultural Heritage (JOCCH)* 6, 2 (2013), 9.
15. J. Hvass, O. Larsen, K. Vendelbo, N. Nilsson, R. Nordahl, and S. Serafin. 2017. Visual realism and presence in a virtual reality game. In *2017 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*. IEEE, 1–4.
16. B. Laugwitz, T. Held, and M. Schrepp. 2008. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*. Springer, 63–76.
17. M. Mortara, C. E. Catalano, F. Bellotti, G. Fiucci, M. Houry-Panchetti, and P. Petridis. 2014. Learning cultural heritage by serious games. *J. of Cultural Heritage* 15, 3 (2014), 318–325.
18. A. Puig, I. Rodríguez, J. Ll. Arcos, J. A. Rodríguez-Aguilar, S. Cebrián, S. Bogdanovych, N. Morera, A. Palomo, and R. Piqué. 2019. Lessons learned from supplementing archaeological museum exhibitions with virtual reality. *Virtual Reality* (29 Jul 2019). DOI: http://dx.doi.org/10.1007/s10055-019-00391-z
19. M. Roussou. 2000. Immersive interactive virtual reality and informal education. In *Proc. of User Interfaces for All: Interactive Learning Environments for Children*. 1–9.
20. G. Schofield, G. Beale, N. Beale, M. Fell, Dawn Hadley, J. Hook, Damian Murphy, Julian R., and L. Thresh. 2018. Viking VR: Designing a virtual reality experience for a museum. In *Proc. of the 2018 Designing Interactive Systems Conference*. ACM, 805–815.
21. M. Seidl, P. Judmaier, F. Baker, C. Chippindale, U. Egger, N. Jax, C. Weis, M. Grubinger, and G. Seidl. 2011. Multi-touch rocks: playing with tangible virtual heritage in the museum-first user tests. In *The 12th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*. 73–76.

22. Blue Byte Software. 1996. HistoryLine: 1914-1918. (1996).