VMUXE
An Approach to User Experience Evaluation for Virtual Museums

Bianca Gockel[1], Holger Graf[1], Alfonsina Pagano[2], Sofia Pescarin[2], and Joakim Eriksson[3]

[1] Fraunhofer Institute for Computer Graphics, Fraunhoferstr. 5, 64283 Darmstadt, Germany
bianca.gockel@gmail.com, holger.graf@igd.fraunhofer.de
[2] CNR-ITABC, Virtual Heritage Lab, Via Salaria, km 29, Monterotondo St., Roma, Italy
alfonsina.pagano@usi.ch, sofia.pescarin@itabc.cnr.it
[3] Lund University, Department of Design Science, Sölvegatan 26, 22100 Lund, Sweden
joakim.eriksson@design.lth.se

Abstract. This paper presents a new approach for the evaluation of User Experience (UX) aspects applied to virtual museums (VM) - VMUXE. A wide percentage of projects and applications for VMs are often "born and buried" in digital labs, without having been experimented and monitored with people. These "prototypes" are the result of experts, technicians, curators, combined together to give birth for multidisciplinary and avant-garde outputs. Earlier attempts to evaluate VM installations failed due to the lack of strategy facing the multidimensional complexity in studying and comparing digital applications in different installations using different devices and metaphors offering different UXs. As a conclusion "communicating" culture through the aid of advanced technology was not a technological issue, but an epistemological one. Setting up a good process of evaluation and analysis is therefore important for establishing next generation virtual museums (NGVM) aiming to reach certain goals such as knowledge exchange, cognitive improvement and heritage communication.

Keywords: User Experience Evaluation, Virtual Museums, (Non-) Instrumental Qualities, Digital Cultural Heritage.

1 Introduction

Today's rapid technological development provides a variety of new possibilities for the cultural heritage sector. In recent years, one major trend became visible as European endeavors aim to digitally master its vast and exclusive cultural resources and learning traditions as source for creativity and inspiration. Michaelis et al. [1] point out, that with the new communication models which are coming along with this evolution the understanding and access to cultural heritage can be enhanced essentially.

A. Marcus (Ed.): DUXU/HCII 2013, Part I, LNCS 8012, pp. 262–272, 2013.
© Springer-Verlag Berlin Heidelberg 2013

One of the trends in digital cultural heritage (DCH) is dedicated to the design and implementation of virtual museums (VMs). The European Commission acknowledges this by funding the Network of Excellence V-MusT.net[1] in order to support this process. Though, no consent on a clear definition for VMs has been found up to now, the V-MusT terminology suggests their nature as being "various digital creations which acquire, conserve, research, communicate and exhibit, in a digital way, the tangible and intangible heritage of humanity and its environment" [2].

As the understanding and learning of our cultural background is inherently aligned to the applied communication models the focus on users' needs, perceptions and interests has to be ensured by a participatory software design. According to Carillo et al. [3] it is already prevalent practice for ICT developments, but not for the implementation of applications in the DCH sector. This needs to be reinforced!

In recent years, some evaluation studies in the DCH domain were conducted. However, those showed a significant fragmentation in their methods, e.g. Carillo et al. [3], Pescarin et al. [4], Micheal et al. [5]. Several authors used a diversity of approaches in different settings (lab vs. in-situ) focusing on certain aspects such as Usability tests using "Cognitive Walkthrough" methods, or considering degree of immersion, behavioral and social aspects, "Smileyometer" and "Again table" of the Fun Toolkit [6] with children to investigate engagement, endurability, and recurrences.

Pujol Tost & Ecounmou [7] predict that lab-based evaluations do not deliver reliable results as they do not take into account the natural environment of exhibitions. On contrary, Rodriguez-Echavarria et al. [8] claim that VM are not well defined as e.g. web application that profit from more standardized hardware and software interfaces and recommend user-based testing at labs. This diversity of opinions is quite predictable as there are specific challenges in the DCH domain.

Thus, the motivation of our paper is to provide a new approach on how we can evaluate the success of a virtual museum aiming for an enhanced user experience. Which are the criteria and quality parameters we can use as reference? What kind of method, if one exists, we should adopt? We present a holistic approach for the design and implementation of an evaluation for VMs – VMUXE. Its analysis will then provide a scaffold of design indicators required for NGVMs.

2 Methodology

2.1 Selection of UX Aspects

Due to the manifold of (non-)instrumental qualities, the selection has been based on "subjective" priorities for interactive VMs. We have been focussing on four selected aspects of UX proposed by the model of Thüring und Mahlke [9]: utility, learnability, efficiency and stimulation.

[1] http://www.v-must.net/ (aof February 2013).

Utility: Requirement analysis that contains the elaboration of user profiles and the contextual task analysis are often not realized in the DCH domain. Therefore, it is vital to examine, if the VM supports all required and desired functions.

Learnability: As users of VMs are often non-experts who only use these applications for example during a visit of a cultural site or even at a physical museum for a limited period of time, it is crucial to ensure an easy learnability of the interface.

Efficiency: As the user continues to use the VM a lack of efficiency might lead to a high level of frustration. Typical objective measures for efficiency are counting clicks or stopping the time for the fulfillment of tasks [10]. However, within this work the subjective perceived efficiency is investigated based on the users' satisfaction.

Stimulation: Although the instrumental needs are fulfilled, excitement and joy which captivate the user are not guaranteed. This term was introduced by Hassenzahl [11] who highlights that individuals strive after reinforcement of their personal development which can be given by providing new, interesting or even exciting content, functionality, presentation or interaction style. Consequently, stimulation as a "driving power" for interaction and learning of new skills and knowledge represents significant quality of VMs.

The focus of this evaluation is thus on the instrumental qualities.

2.2 Selection of Evaluation Methods

For the elevation of data a combination of two methods was chosen, as to achieve a holistic picture highlighted by Lehn et al. [12]. A written survey in form of a paper-based questionnaire was designed with the aim of collecting quantitative data that should be used to identify tendencies for the levels of maturity concerning the different selected UX aspects that are statistical significant. The questionnaire follows ensured structure:

1. *Part A - Demographic Data* that collects essential information about the users' characteristics (gender, age, origin, occupation and ICT knowledge)
2. *Part B - Feedback* that contains:
   (a) an extract from the User Experience Questionnaire (UEQ) [13] that summarizes the results of 12 pairs of contrasting attributes with a seven point scale to the terms perspicuity, efficiency and stimulation
   (b) several multiple choice questions to deliver concrete statements (see tab. 1)

As a completion, an *interview guideline* based on a questionnaire was developed in order to identify the motivation behind the users' reactions and deduce some suggestions for improvements for the concept of a VM, with a focus on qualitative data and is structured as follows:

Table 1. Written Survey: Part A - Feedback: Multiple choice questions

UX Aspect	*Question*	*Answer Category*
Utility	**Q1**: What do you think about the functions of the virtual museum?	**AC1**: All functions are useful and interesting to me **AC2**: The virtual museum contains some interesting functions, but also some unnecessary functions, namely:_____ **AC3**: No functions are useful and interesting for me
	Q2: Was every function available that you wish to have?	**AC1**: Yes **AC2**: No, I would like to add following function:_____
Learnability	**Q3**: What do you think about the handling of the virtual museum?	**AC1**: It was hard to use the virtual museum, so I couldn't focus on the content. **AC2**: At first I had to fully concentrate to use the virtual museum, but after some time I could focus on the content. **AC3**: From the start I had no problems to use the virtual museum and could fully enjoy it.
Efficiency	**Q4**: What do you think about the time it took to use it?	**AC1**: I am satisfied with the time that took me to reach my goals and I could fully enjoy the content. **AC2**: Sometimes it took too much time, because the system reacted slowly. **AC3**: Sometimes it took too much time, because I had to perform unnecessary and/or inconvenient steps. **AC4**: It took me so much time, that I completely lost focus of the content.
Stimulation	**Q5**: What do you think about your experience with the virtual museum?	**AC1**: I experienced something totally new for me and got new insights. **AC2**: Experienced something I already know, but I could still enjoy **AC3**: Experienced something already known a lot and boring.

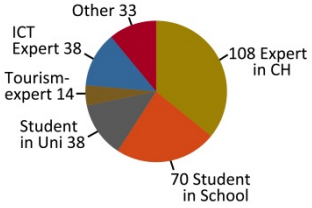
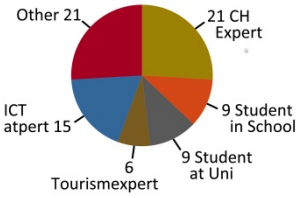
1. *Part A - Demographic Questionnaire* that is equal to the one of the written survey
2. *Part B - Demonstrator Use* that instructs the participant to accomplish tasks adapted to each VM while applying "Thinking Aloud" observed by the evaluator
3. *Part C - Final Questions* that is similarly structured as the multiple choice questions of the written survey, but more broken-down and opened-up designed

3 Results

3.1 Practical Implementation

We conducted the evaluation during the Archeovirtual 2012[2] which offers showcases of various VMs from all over the world. 24 scientifically reviewed projects were allocated to 5 technological areas: Mobile, Touch, Computer Animation, Emerging Technologies, Natural Interaction and Desktop Virtual Reality. The written survey was applied to several presented 18 *interactive* VMs. 10 most innovative and interesting projects were chosen for a deeper evaluation in face-to-face interviews. Table 2 shows the total numbers that could be reached for the whole evaluation study.

Table 2. User Profile of the Archeovirtual 2012

		Written Survey (N = 301)	*Interview (N = 81)*
Gender	Female	157	34
	Male	144	47
Age	Range	11 – 70 years	7 – 63 years
Occupation	Composition		

Due to the scope of this paper, we decided to present the results of three VMs for a special technological domain that mirrors the results for the overall domain. The results for desktop based VR can be found in [14]. Several quantitative results were mapped onto a scale between -3 and 3. The neutral interval contains values between -0.8 and 0.8. Extreme values at the end of the scale either smaller than -2 or bigger than 2 are rated as “very poor” or “very good”.

3.2 Area Mobile: Chartreuse Numérique

This VM offers an interactive 3D real-time visit for the iPad of the Charterhouse’s church [3] of Villeneuve-lès-Avignon (Gard, France) based on panoramic photos obtained by superposition between virtual images (extracted from a 3D model) and images of the real space. A tap on an interactive map chooses a location inside the

[2] Archeovirtual – Embedded within Mediterranean Expo for Archaeological Tourism, in 2012: 15th-18th of November, Paestum, Italy, `www.archeovirtual.it` (aof February 2013).

[3] Project by Prof. Livio De Luca, MAP laboratory, UMR CNRS/MCC 3495, ENSA Marseille, France - `http://www.chartreuse.org/16/337/espace-3d-chartreuse` (aof February 2013).

church. By using a slider the panoramic picture is increasingly overlaid by the 3D reconstruction. Swiping on the touchscreen allows changing the perspective and tapping paintings zoom the user in. Further descriptive information about the paintings can be accessed by tapping on an icon located at the lower right corner (see fig. 1). 22 written survey and 7 interviews were carried out for the VM. Table 3 shows the quantitative data that could be achieved for perspicuity, efficiency and stimulation.

For *learnability* only an average result could be achieved. It was enhanced by simple map navigation, but also lowered by two main problems: First the 3D reconstruction could not be related to a specific historical time as the color concept of the interactive map was not clear and the time slider was not labeled. Second, the accessibility of the paintings is for the users not transparent as these hotspots are not specifically highlighted compared to other non-interactive parts. For *efficiency* a good result could be obtained with a tendency to very good as only 21% within the written survey detected inefficiencies. The *stimulation* values range between average and good, as the combination of content and functions appeared new.

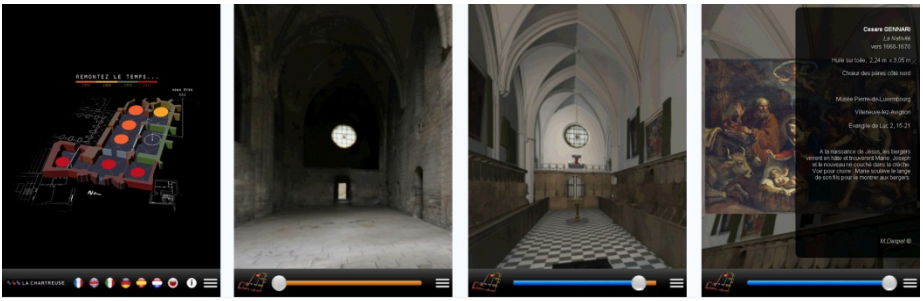
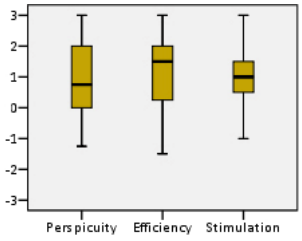


Fig. 1. Chartreuse Numérique - Left: Interactive map as starting point, Mid: Location inside the church at present state or at past state, Right: Description about selected paintings

Table 3. Chartreuse Numérique - Scales

		Perspicuity	*Efficiency*	*Stimulation*
N	Valid	21	21	21
	Missing	1	1	1
Median		.75	1.50	1.00
Std. Deviation		1.19784	1.26220	.89642
Rating		Neutral	Good	Good



The *utility* of this VM was marked as very good. Within the written survey 91% of participants rated all available functions as useful. Only the transition from present to past by changing the opacity was criticized. A wish list does provide suggestions for further developments: More descriptions and storytelling, move from room to room without going back to the map, links between paintings, audio comments, more hotspots inside the church, e.g. windows.

3.3 Area Emerging Technologies: Bravo – BRAin Virtual Operator

The emerging technology area was dedicated to Brain-Computer Interfaces (BCI). BRAVO (BRAin Virtual Operator)[4] represents a user-centered and portable solution that applies BCI to e-learning. It is a tool for reading customisable eLearning courses and multimedia content using a BCI acquiring information about attention and relaxation levels (see fig.2).

While for the scale *perspicuity* a good level was achieved, *learnability* reached only an average level with a predisposition to good as to a large extent it required some time to understand how to use the VM (see table 4). *Efficiency* reached a good level with a tendency to very good as only 11% mentioned that the system reacted sometimes too slow.

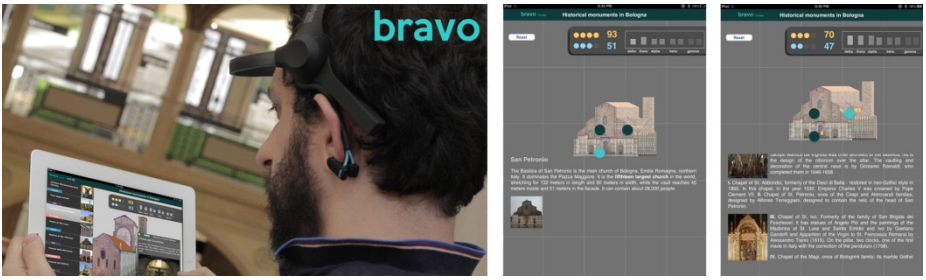
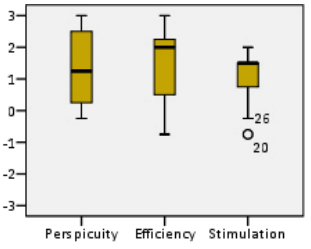


Fig. 2. Bravo - Left: Setting, Right: Screenshots of adapted content

Table 4. Bravo - Scales

		Perspicuity	*Efficiency*	*Stimulation*
N	Valid	26	25	25
	Missing	2	3	3
Median		1.3750	2.0000	1.5000
Std. Deviation		1.19422	1.11822	.83229
Rating		Good	Good	Good



In addition, for *stimulation* a good result was achieved as it was for a majority of 79% a totally new experience to interact with such a VM. Nobody disliked the VM or felt bored. For *utility*, 96% of the participants rated that all available functions are useful and interesting to them. Suggestions for enhancements: A higher variety of media types and a more reliable auditing system.

[4] Project by Dr. Marco Marchesi, University of Bologna, DEIS Dept. of Electronics, Italy.

3.4 Area Natural Interaction: Rediscovering Vrouw Maria

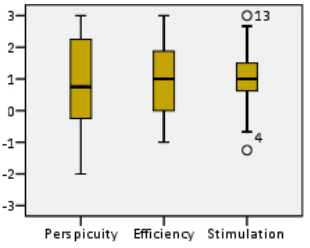
Rediscovering Vrouw Maria[5] is an interactive, real-time, virtual reality simulation about the Vrouw Maria shipwreck that uses stereoscopic projection. The interaction is gesture based tracked by the Microsoft's Kinect Sensor. The user can navigate freely through the landscape and inspect a selection of the remains of the wreck displayed that can be chosen by moving through the particular info-icon (see fig.3).

Only an average result was obtained for *learnability* as it was decreased by initial problems to understand which gesture leads to which system reaction. After a short learning phase the gestures could easily be applied. For *efficiency* different results were obtained. While in the interview all participants were satisfied, within the written survey 32% detected inefficiencies. For *stimulation* a good result was reached due to the immersive environment that allows exploring the wreck in a new way. For *utility* all available functions were appreciated by 92% who participated at the written survey. Only one participant rated the selection of information by moving through an icon as not useful (see table 5). Enhancements: More objects/remains to investigate under the sea and audio, "swim" gestures, more contextual information.

Fig. 3. Rediscovering Vrouw Maria - Left: Screnshot, Right: Person interacting with the VM by pointing with the hand

Table 5. Vrouw Maria - Scales

		Perspicuity	*Efficiency*	*Stimulation*
N	Valid	24	24	24
	Missing	0	0	0
Median		.75	1.00	1.00
Std. Deviation		1.50843	1.18793	1.00842
Rating		Neutral	Good	Good



[5] Project by Prof. Lily Diaz-Kommonen, Aalto University School of Arts Design and Architecture, Aalto, Finland - http://sysrep.aalto.fi/vrouwmaria/ (aof February 2013).

4 Discussion

As all VMs presented at Archeovirtual 2012 faced a technological-based categorization, the subsequent discussion highlights general tendencies for the thematic groups and reveals weaknesses and strengths of VMUXE.

Mobile Area: VMUXE mainly revealed issues on the visibility of system's elements and content redundancy. This is obviously reflected in the usability of applications, leaving users disoriented by misleading information and confusing interfaces. The same visual and textual indicators (i.e. hotspots' accessibility) suffered from consistency and users' orientation, e.g. the paintings' accessibility within Chartréuse Numerique using highlights for non-/interactive parts of the church. Apart from this, VMs of this category were appreciated as involving and well structured.

Emerging Technologies: required a constant maintenance by the particular developers on site because of the necessity of a preliminary introduction to users (fitting of the BCI headset, explanation on the interaction). Here, the categories of learnability and efficiency resulted to be irrelevant for our study as the time needed to reach certain goals is planned to be a result of the applications not a parameter to be investigated. Also the easiness of the system turned to be a subjective datum, as user alone could not grab if he succeeded unless the developer explained it to him. In general, the two presented VMs are more suitable for visibility and aesthetical evaluations since interface and visual elements resulted the most interesting items highlighted by users (i.e. more multimedia functions, changes in icons, customization of characters).

Natural Interaction Area: all of the presented VMs were appreciated for their immersiveness, the natural atmosphere recreated, and good feedback of the system. Nevertheless, some efforts were suggested wrt. content redundancy (i.e. too much text in the Wreck objects' explanations), visibility of interface elements (i.e. orientation map of Vrouw Maria which indicated the position of the user in relation to the wreck were not clearly identifiable), and mapping between users' controls and system's effects. Again for this category, the presence of developers on site was essential.

In the end, speaking about UX aspects and what can positively influence their values, we can affirm that *utility* is founded on the consistency of interface's elements, linearity and conciseness of contents and a good mapping between users' control and system's effects; *learnability* sees connections with a good visibility of navigation indicators, affordance of visual and textual information and the consistent nature of contents delivered to users; *efficiency* takes advantages of the rapid feedback given by the system to users and the multiplicity of functions offered, yet, its questionable to further inspect this as users should be granted some time to explore a VM; *stimulation* finds help by the integration of media, which provides pleasant aesthetical features, the natural setting presented by the whole environment and the easiness of interaction.

5 Conclusion

One can certainly expect that NGVMs appearing in forthcoming years will have a wide variation in interaction techniques, as well as in the level of ambition regarding information/education. Thus, it is of course a challenge to use a uniform evaluation method for all these different applications.

With VMUXE, we propose an evaluation method that addresses both instrumental and non-instrumental qualities, and collects both quantitative as well as qualitative data. The case studies performed at Archeovirtual, showed that the method works well and is efficient, especially considering the special circumstances in field-testing on museum visitors, as opposed to how a controlled environment study could have been performed. When evaluating with VMUXE, there will naturally be some applications where certain questions may feel irrelevant to ask people, and some UX-aspects may feel of less importance, such as efficiency.

However, the generative approach provides the benefit to gradually build up some benchmark values, based on a larger number of investigations. It would then become easier to validate a new VM's status by comparing results with existing benchmarks.

Acknowledgements. The work is funded by the European Commission through the European Community's Seventh Framework Programme (FP7 2007/2013) under grant agreement 270404 – "V-Must.net".

References

1. Michaelis, N., Jung, Y., Behr, J.: Virtual Heritage to go. In: Proceedings of the 17th International Conference on 3D Web Technology, pp. 113–116. ACM, New York (2012)
2. Consortium of the V-MusT.net: Terminology, Definitions and Type for Virtual Museums. In: Deliverable Report: V-MusT.net – D 2.Version 3 (2013)
3. Carillo, E., Rodriguez-Echavarria, K., Heath, R., Arnold, D.: Effective user-centred design for the Vindolanda interactive multimedia application. In: Digital Heritage, Proceedings of the 14th International Conference on Virtual Systems and Multimedia, Short Papers, VSMM 2008, pp. 337–342. Archeolingua, Limassol (2008)
4. Pescarin, S., Pagano, A., Wallergård, H.W., Ray, C.: Archeovirtual 2011: An Evaluation Approach to Virtual Museums. In: Proceedings of the 18th Conference on Virtual Systems and Multimedia, VSMM 2012, pp. 25–32. IEEE Press, Milan (2012)
5. Michael, D., Pelekanos, N., Chrysanthou, I., Zaharias, P., Hadjigavriel, L.L., Chrysanthou, Y.: Comparative Study of Interactive Systems in a Museum. In: Ioannides, M., Fellner, D., Georgopoulos, A., Hadjimitsis, D.G. (eds.) EuroMed 2010. LNCS, vol. 6436, pp. 250–261. Springer, Heidelberg (2010)
6. Read, J., MacFarlane, S., Casey, C.: Endurability, Engagement and Expectations: Measuring Children's Fun. In: Interaction Design and Children, pp. 189–198. Shaker Publishing, Eindhoven (2002)
7. Pujol Tost, L., Economou, M.: Evaluating the Social Context of ICT Applications in Museum Exhibitions. In: Proceedings of the 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage, VAST 2006, pp. 219–228. Eurographics Association, Nicosia (2006)

8. Rodriguez-Echavarria, K., Morris, D., Moore, C., Arnold, D., Glauert, J., Jennings, V.: Developing Effective Interfaces for Cultural Heritage 3D Immersive Environments. In: Proceedings of the 8th International Symposium on Virtual Reality, Archaeology and Cultural Heritage, VAST 2007, pp. 93–99. Eurographics Association, Brighton (2007)
9. Thüring, M., Mahlke, S.: Usability, aesthetics, and emotion in human-technology interaction. International Journal of Psychology 42, 253–264 (2007)
10. Nielsen, J.: Usability Engineering. Morgan Kaufmann, San Francisco (1994)
11. Hassenzahl, M.: The thing and I: understanding the relationship between user and product. In: Blythe, M.A., Overbeeke, K., Monk, A.F. (eds.) Funology: From Usability to Enjoyment, pp. 31–42. Kluwer Academic Publisher, Dordrecht (2003)
12. Vom Lehn, D., Heath, C., Hindmarsh, J.: Video based field studies in museums and galleries. Visitor Studies Today! 5(3), 15–23 (2002)
13. Laugwitz, B., Held, T., Schrepp, M.: Construction and Evaluation of a User Experience Questionnaire. In: Holzinger, A. (ed.) USAB 2008. LNCS, vol. 5298, pp. 63–76. Springer, Heidelberg (2008)
14. Baldassari, G.L., Demetrescu, E., Pescarin, S., Eriksson, J., Graf, H.: "Villa di Livia": A case study for the devolution of large scale interactive "in-site" to "on-line" application. In: Proceedings of the HCI International 2013. LNCS. Springer, Heidelberg (2013)