Open the microphone, please! Conversational UX Evaluation in Virtual Reality

I. Rodríguez
inmarodriguez@ub.edu
Departament de Matemàtiques i Informàtica.
Facultat de Matemàtiques i Informàtica, UBICS
UB Institute of Complex Systems, UBICS
University of Barcelona

A. Puig
annapuig@ub.edu
Departament de Matemàtiques i Informàtica.
Facultat de Matemàtiques i Informàtica
Institut de Matemàtica de la UB, IMUB
University of Barcelona

ABSTRACT

This paper proposes the use of conversational interactions for gathering feedback from users in Virtual Reality (VR) evaluation studies. A conversational interaction allows the user to communicate with the system using natural language in the form of text, voice or both. This interaction is facilitated by what are known as conversational agents (CA), which engage in a conversation with the user. In contrast to gathering user feedback once the experience has finished, these agents, either embodied or not, are in charge of administering the user post-tasks and post-study questionnaires, which are carried out inside the VR environment. In this paper we conceptualise conversational agents in the context of UXE (User eXperience Evaluation) in VR, and analyse key design elements to be taken into consideration when designing them. We hope our discussion encourages others to study in-world evaluation in VR.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; • **Computing methodologies** → **Virtual reality**.

1 INTRODUCTION

The availability of affordable Head-Mounted Displays (HDM) and Software Development Kits (SDK) has extended the use of Virtual Reality (VR) in different fields, such as the automotive, healthcare, education and entertainment industries, among others [21][10]. It is therefore increasingly necessary to conduct user studies to validate and improve developments in VR.

VR experience evaluation can be performed in two-ways: in-world (while the user is in the virtual environment) and out-world (once the experience is over) [5]. In-world evaluation can be done implicitly and explicitly, through logs running in the background and by asking the user to fill in questionnaires, respectively.

Regarding questionnaires, there are two modalities: post-task and post-study [24]. In the context of VR experiences, it is clear that the post-task modality can hardly be done out-world and is, therefore, almost forcibly carried out in-world. In contrast, the post-study modality can be performed either way.

In-world questionnaires have the advantage of not interrupting the immersive experience [25][17][5]. Nevertheless, it runs into the barrier posed by VR input methods. For example, questionnaires designed as point and click planar surfaces in the 3D virtual world can be cumbersome and unnatural. Conversational interfaces are currently gaining popularity as an alternative, and they provide the user with a natural interaction using text and voice [9] [14]. Although the limitation of input methods for HMDs may have been a handicap to implement conversational interfaces in the past, this is not the case today thanks to commercial devices such as 6DOF controllers and microphones.

In this panorama of powerful and affordable VR devices, we advocate for exploring new methods of UX evaluation (UXE) in VR. Concretely, we propose the evaluation of VR experiences through conversational (post-task and post-study) questionnaires performed in-world, i.e while the users are living the immersive experience. This is, as far as we know, the first proposal towards the introduction of conversational interactions when evaluating VR applications. In this paper, we discuss the opportunities and challenges arising from our proposal, laying out both our theoretical and practical thoughts, while hoping it will encourage further theoretical and experimental studies on the topic.

2 RELATED WORK

Classical methods for evaluating UX employ a wide spectrum of measures, ranging from objective (heart rate, sweating, eye tracking) to subjective ones (questionnaires) [4]. Similar measures have been applied to VR applications, with the addition of platform-dependent metrics such as gaze and 3D position-orientation tracking [26] [19].

In the following, we focus on research works concerned with gathering users' opinions and feelings using questionnaires. First, we present research that focuses on the use of questionnaires for the evaluation of UX experience in VR. Then, we refer to studies that use conversational interactions for administering surveys to users.

The literature reports several studies that use in-world questionnaires in VR, ranging from a single question with multiple responses operated by a virtual pointer or a user tracked hand [13] [29] to sliders on a 0 to 10 scale [6] [25]. A recent study [5] compared the use of in-world and out-world questionnaires, dealing with different dimensions such as enjoyment, efficiency (completion time) and, in general, usability. In-world showed better results in terms of enjoyment but usability and efficiency scored higher in out-world questionnaires. The explanation for this can be found in users' greater familiarity with mouse and keyboard interactions than with their virtual counterparts. While these results favour the use of in-world questionnaires, the authors suggest more research work is needed to assess their impact on reliability.

Some initial steps have also been taken recently with regard to the use of conversational interactions to perform surveys and questionnaires. Celino [8] proposed a conversational survey tool

Table 1: Classification parameters of conversational agents

Classification parameter	Value of parameter	Parameter in UXE VR
Aspect	non-embodied, both embodied	non-embodied, embodied
Knowledge domain	open-domain, closed-domain	closed domain
Service provided	interpersonal, intrapersonal	both interpersonal and intrapersonal
Goals	chat-based, task-based	task-based
Method	rule-based, retrieval-based, generative	rule-based

that administer questionnaires through a chat-like interface. Experimental results showed that users expressed a preference for the conversational interface. Along the same lines, Xiao [31] used an AI based chatbot to conduct open-question surveys and interviews. The chatbot generated a higher level of participant engagement and higher quality responses than the classical survey, as well as outperforming it in terms of trustworthiness [3]. Our research relies on previous work carried out along both lines of research - assessing VR experience in-world through questionnaires and conversational interactions for administering questionnaires - and proposes the use of conversational interactions for facilitating UXE in virtual reality applications.

3 IN-WORLD CONVERSATIONAL INTERACTIONS

In this section, we first conceptualize conversational agents in the context of UXE in VR and then we discuss some design elements to consider when designing them.

3.1 Conversational agents for UXE

Conversational agents can be classified according to features such as aspect (embodied, non-embodied), interaction mode (text, voice based), knowledge (open-domain, specific-domain), method (rule-based, AI) and application (task-oriented, non-task oriented) [12][1]. We take agent features from previous classification proposals and provide a rationale for agent parameters in UXE VR. Table 1 shows the different parameters - aspect, domain knowledge, service provided, goals, method - that are presented next, and, in the last column, indicates the value of the parameter that best suits UXE VR.

Embodied conversational agents (ECA) are anthropomorphic representations of the system [7]. They are adequate for those VR applications which include, by design, narratives and characters that drive the narrative or act as companions for the players. Nevertheless, conversational interactions can be also designed featuring non-embodied agents, as is the case of VR desktop applications, for example.

Agents' knowledge may be general or specific, i.e. there are agents who engage in conversations on general topics (open-domain) and agents who know about a particular domain (closed-domain). Conversational agents for UXE VR clearly belong to a concrete domain and their knowledge would be restricted to specific questions regarding the user experience and users' answers that are designed to elicit feedback on their experience.

Regarding the service provided, interpersonal agents actuate as facilitators for reaching a goal, and intrapersonal agents actuate as a companion with a emotion based relationship. For UXE VR, we envisage both possibilities, depending on the VR application. For example, a desktop-based VR aplication may be better adapted to interpersonal service because the agent is specifically designed for the purpose of the evaluation. Conversely, VR gamified applications and games fit better with an intrapersonal service, because the agent is also designed to elicit empathy and emotional experiences [2][27]. An agent who is in charge of facilitating the evaluation of a VR experience is designed to perform a specific task (task-based), and its purpose is not, therefore, to give information or chat about a particular topic (chat-based).

Finally, there are several methods for processing user inputs and generating corresponding responses. In order of increasing complexity these are rule-based methods (in which the users' input sentences match one of the expected inputs, and the response is a sentence based on a fixed predefined set)[23], retrieval-based methods (which gather candidate responses from available services) [30] [11] and generative (responses are generated by AI systems) [18] [22] . Given the simplicity of the interactions needed for gathering user feedback through in-world questionnaires, we consider rule-based systems (without creating answers as generative systems do) to be adequate for designing and implementing UXE VR agents.

3.2 Design considerations

When designing a conversational experience there are also a number of design elements to consider, some of which are general enough to be taken into account for a variety of platforms (desktop, mobile, etc.), while others are more directly related to 3D virtual environments like VR. Figure 2 shows general design elements in blue - dialogue flow, context, language and tone, user privacy and trust, and dealing with failures - and those related to VR are shown in green - 3D avatar and mix of input methods. We elaborate on these elements in the following.

A *dialogue* consists of a sequence of questions and answers connected linearly or forming branches depending on the user's choices. That is, the dialogue can be a simple question-answer sequence or a more complex interaction defined by an (underlying) dialogue flow. In the context of a UXE in VR we think that dialogues can be simple non-linear ones. For example, Figure 1 shows a short dialogue with three branches: the agent asks the user "Did you feel immersed while doing the task?", if the user response is affirmative ("Yes", "Of course", "Sure I did"...), the agent reply can be "Wow, these are good news!". But if the response is "No", the agent reply can be "What a pity. Can you tell me why you did not feel immersed?". If the user gives a response, the agent thanks the user for the response, if not, the agent finishes the conversation by saying "All right, bye". Other user responses, such as "Not sure" and "I don't know" can be answered with "Ok, thanks, don't worry about that!".

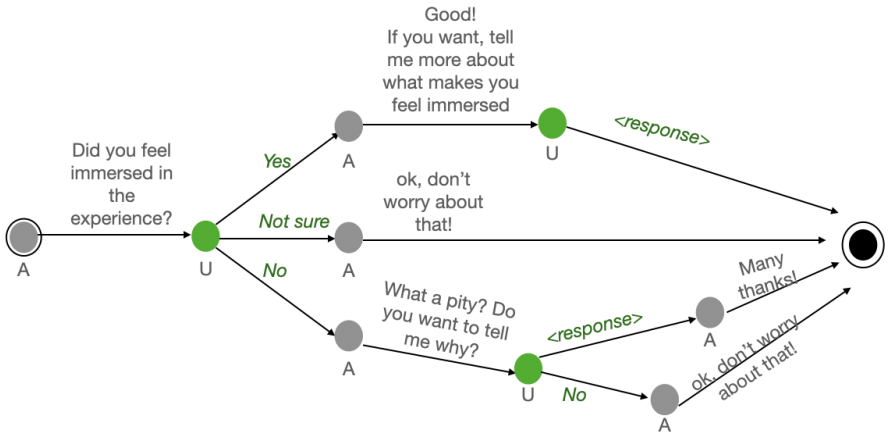

Figure 1: User(U)-agent(A) dialog for assessing UX experience.

Language and tone refer to the use of formal and informal language in conversational interactions. Previous research has discussed the importance of selecting the correct register and tone [8], since the target users and the type of VR application determine the register and tone of the conversation.

Moreover, it is important to maintain the *context* during a conversation, such as the agent recognising a change of topic during the conversation and remembering topics that have refereed to earlier in the conversation. This may have an impact on users' perception of the agents' credibility and usefulness. For example, if the user refers to previous opinions, the agent should be able to recognise that in order to give an appropriate response [16].

Related to *user privacy and trust*, conversational agent design should establish clear communications with the user regarding data privacy and data management, since trust is important for the success of human-agent interaction. Additionally, chatbots have demonstrated to allow users' self-disclosure while securing their anonymity [28]. This act of revealing information can be an opportunity for UXE, since the goal is to gather reliable data regarding users' opinions and perceptions of their experience.

Agents often fail to fulfil users' expectations because of their limited understanding capabilities [15], which means that it is crucial to *deal with failures* related to natural language processing. Moreover, other failures, such as technical issues (e.g. failed internet connection, speech-to-text system malfunctions), should be also managed by the system.

Regarding the elements that are closely related to UXE in VR, a *3D avatar* is a character that can introduce the application to users and guide them through the experience. The avatar could also be used as a facilitator for the conversational in-world questionnaire, thereby allowing for a seamless experience for the user. The use of embodied conversational agents favour their integration into the narrative, providing the agent with a role and a goal. This role may be to act purely as a test moderator asking the user about several aspects of the experience (this can be done in different parts of the designed experience) or as a companion interested in the overall experience lived by the user.

It is proven that users prefer a mix of *input methods*, i.e., offering the user several possibilities to provide their input for the system. Taking into account the fact that interactive 2D elements (e.g. menus) can be cumbersome to manipulate in VR, we believe that microphones can provide users with a more natural way to answer questions regarding their UX experience within the VR environment. Indeed, the use of microphones could at some point violate the privacy and anonymity of users, but it can be preserved adding sound masks to anonymize speech data before being sent to the storage system [20].

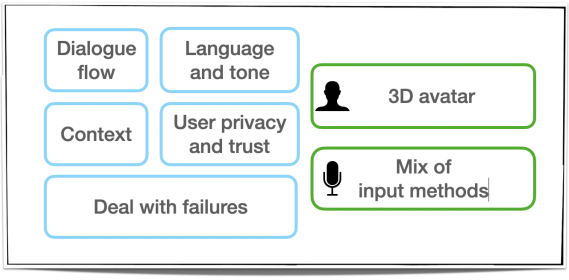

Figure 2: Design elements in conversational UXE. Blue (general) Green (VR-related)

To conclude, the conversational approach is an interaction style that can be natural and efficient, since it allows the user to respond by voice to UX related questions. Additionally, users do not tend to forget what happened to them and how they felt in an in-world conversational questionnaire, and the reporting of their feelings inside the virtual environment may provide better quality of data. Nevertheless, further experimental research is needed in order to provide evidence regarding this issue. There are also a number of limitations to this approach, such as the design of the conversational system, which is an additional effort that should done by the designer. Moreover, when the system includes conversational

Figure 3: Conversational interaction in a VR game.

interactions for purposes other than evaluation, it should be always clear for the user what is the purpose of the conversation. Therefore, the agent should take the context into consideration more than ever, not only because of its past history in the conversation but also in order to make a clear distinction between talking about the experience and talking during the experience.

4 CONCLUSIONS

This paper proposes the use of in-world questionnaires to gather users opinions and likes in virtual reality applications. Specifically, a conversational interaction style may enhance the efficiency and naturalness of the evaluation process. As on-going work, we are integrating a conversational agent in a serious game that introduces programming concepts and computational thinking to tweens. This research is a starting point to be followed by theoretical and technological developments as well as experimental studies to compare it with alternative interaction styles. Further studies are needed to evaluate the impact of conversational in-world questionnaires on both UX and questionnaire reliability.

ACKNOWLEDGMENTS

To MISMIS-Language (PGC2018- 096212-B-C33), CI-SUSTAIN (PID2019-104156GB-I00) and Accio COMRDI18-1-0010.

REFERENCES

[1] Eleni Adamopoulou and Lefteris Moussiades. 2020. An overview of chatbot technology. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer, 373–383.

[2] Imtiaj Ahmed, Ville Harjunen, Giulio Jacucci, Niklas Ravaja, and Michiel M Spapé. 2016. Total immersion: Designing for affective symbiosis in a virtual reality game with haptics, biosensors, and emotive agents. In *International Workshop on Symbiotic Interaction*. Springer, Cham, 23–37.

[3] Fatema Akbar, Ted Grover, Gloria Mark, and Michelle X Zhou. 2018. The Effects of Virtual Agents' Characteristics on User Impressions and Language Use. In *Proc. 23rd International Conference on Intelligent User Interfaces Companion*. 1–2.

[4] William Albert and Thomas Tullis. 2013. *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Newnes.

[5] Dmitry Alexandrovsky, Susanne Putze, Michael Bonfert, Sebastian Höffner, Pitt Michelmann, Dirk Wenig, Rainer Malaka, and Jan David Smeddinck. 2020. Examining design choices of questionnaires in vr user studies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–21.

[6] Zekun Cao, Jason Jerald, and Regis Kopper. 2018. Visually-induced motion sickness reduction via static and dynamic rest frames. In *2018 IEEE conference on virtual reality and 3D user interfaces (VR)*. IEEE, 105–112.

[7] Justine Cassell. 2001. Embodied conversational agents: representation and intelligence in user interfaces. *AI magazine* 22, 4 (2001), 67–67.

[8] Irene Celino and Gloria Re Calegari. 2020. Submitting surveys via a conversational interface: an evaluation of user acceptance and approach effectiveness. *International Journal of Human-Computer Studies* 139 (2020), 102410.

[9] Leigh Clark, Philip Doyle, Diego Garaialde, Emer Gilmartin, Stephan Schlgl, Jens Edlund, Matthew Aylett, Jo£o Cabral, Cosmin Munteanu, Justin Edwards, and Benjamin R Cowan. 2019. The State of Speech in HCI: Trends, Themes and Challenges. *Interacting with Computers* 31, 4 (09 2019), 349–371.

[10] Caroline Fertleman, Phoebe Aubugeau-Williams, C. Sher, Ai-Nee Lim, S. Lumley, S. Delacroix, and Xueni Pan. 2018. A discussion of virtual reality as a new tool for training healthcare professionals. *Frontiers in public health* 6 (2018), 44.

[11] Ho Thao Hien, Pham-Nguyen Cuong, Le Nguyen Hoai Nam, Ho Le Thi Kim Nhung, and Le Dinh Thang. 2018. Intelligent assistants in higher-education environments: the FIT-EBot, a chatbot for administrative and learning support. In *Proceedings of the ninth international symposium on information and communication technology*. 69–76.

[12] Shafquat Hussain, Omid Ameri Sianaki, and Nedal Ababneh. 2019. A survey on conversational agents/chatbots classification and design techniques. In *Workshops of the International Conference on Advanced Information Networking and Applications*. Springer, 946–956.

[13] HyeongYeop Kang, Geonsun Lee, Seongsu Kwon, Ohung Kwon, Seongpil Kim, and JungHyun Han. 2018. Flotation Simulation in a Cable-driven Virtual Environment–A Study with Parasailing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–11.

[14] Raina Langevin, Ross Lordon, Thi Avrahmi, Benjamin Cowan, Tad Hirsch, and Gary Hsieh. 2021. Heuristic Evaluation of Conversational Agents. (2021).

[15] Ewa Luger and Abigail Sellen. [n.d.]. " Like Having a Really Bad PA" The Gulf between User Expectation and Experience of Conversational Agents. In *CHI, pages=5286–5297, year=2016*.

[16] Vinayak Mathur and Arpit Singh. 2018. The rapidly changing landscape of conversational agents. *arXiv preprint arXiv:1803.08419* (2018).

[17] Sebastian Oberdörfer, David Heidrich, and Marc Erich Latoschik. 2019. Usability of gamified knowledge learning in VR and desktop-3D. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

[18] Mayur Patidar, Puneet Agarwal, Lovekesh Vig, and Gautam Shroff. 2017. Correcting Linguistic Training Bias in an FAQ-bot using LSTM-VAE. In *DMNLP Workshop of ECML-PKDD*.

[19] A. Puig, I. Rodríguez, J. L. Arcos, J.A. Rodríguez-Aguilar, S. Cebrián, A. Bogdanovych, N. Morera, A. Palomo, and R. Piqué. 2020. Lessons learned from supplementing archaeological museum exhibitions with virtual reality. *Virtual Reality* 24, 2 (2020), 343–358.

[20] Jianwei Qian, Haohua Du, Jiahui Hou, Linlin Chen, Taeho Jung, and Xiang-Yang Li. 2018. Hidebehind: Enjoy voice input with voiceprint unclonability and anonymity. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*. 82–94.

[21] Washington X Quevedo, Jorge S Sánchez, Oscar Arteaga, Marcelo Álvarez, Víctor D Zambrano, Carlos R Sánchez, and Víctor H Andaluz. 2017. Virtual reality system for training in automotive mechanics. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*. Springer, 185–198.

[22] AM Rahman, Abdullah Al Mamun, and Alma Islam. 2017. Programming challenges of chatbot: Current and future prospective. In *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*. IEEE, 75–78.

[23] Md Shahriare Satu, Md Hasnat Parvez, et al. 2015. Review of integrated applications with aiml based chatbot. In *2015 International Conference on Computer and Information Engineering (ICCIE)*. IEEE, 87–90.

[24] Jeff Sauro and James R Lewis. 2016. *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann.

[25] Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. 2019. Using presence questionnaires in virtual reality. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–12.

[26] Alistair G Sutcliffe and K Deol Kaur. 2000. Evaluating the usability of virtual reality user interfaces. *Behaviour & Information Technology* 19, 6 (2000), 415–426.

[27] D. Tellols, M. Lopez-Sanchez, I. Rodríguez, P. Almajano, and A. Puig. 2020. Enhancing sentient embodied conversational agents with machine learning. *Pattern Recognition Letters* 129 (2020), 317–323.

[28] Chris van der Lee, E. Croes, J. de Wit, and M. Antheunis. 2019. Digital confessions: Exploring the role of chatbots in self-disclosure. *Evaluation* 7 (2019), 21.

[29] Carolin Wienrich, Kristina Schindler, N Döllinqer, Simon Kock, and Ole Traupe. 2018. Social presence and cooperation in large-scale multi-user virtual reality-the relevance of social interdependence for location-based environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 207–214.

[30] Yu Wu, Wei Wu, Chen Xing, Can Xu, Zhoujun Li, and Ming Zhou. 2019. A Sequential Matching Framework for Multi-Turn Response Selection in Retrieval-Based Chatbots. *Computational Linguistics* 45, 1 (March 2019), 163–197.

[31] Ziang Xiao, Michelle X Zhou, Q Vera Liao, Gloria Mark, Changyan Chi, Wenxi Chen, and Huahai Yang. 2020. Tell Me About Yourself: Using an AI-Powered Chatbot to Conduct Conversational Surveys with Open-ended Questions. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 3 (2020), 1–37.