Proposition and Validation of a Questionnaire to Measure the User Experience in Immersive Virtual Environments

Katy Tcha-Tokey [1], Olivier Christmann [1], Emilie Loup-Escande [2] and Simon Richir [1]

[1] LAMPA EA 1427, Arts et Métiers Paris Tech, 2 bd du Ronceray, 49000 ANGERS, France
[2] CRP-CPO EA 7273, Université de Picardie Jules Verne, Chemin du Thil, 80000 AMIENS, France

Abstract - There are increasing new advances in virtual reality technologies as well as a rise in immersive virtual environments research and user experience research. Within this framework, we decided to address the overall user experience in immersive virtual environments. Indeed, in our point of view, this topic is not fully dealt with in the scientific literature, neither in terms of user experience components nor in terms of user experience measurement methods. It is in this context that we conducted a study aiming at proposing and validating a unified questionnaire on user experience in immersive virtual environment(IVEQ). Our questionnaire contains 10 subscales measuring presence, engagement, immersion, flow, usability, skill, emotion, experience consequence, judgement and technology adoption. The construction of our questionnaire was based on existing ones. It was tested on 116 participants after they use the edutainment virtual environment "Think and Shoot". Results show that 9 out of 10 subscales and 68 out of 87 items are reliable as demonstrated by an internal consistency analysis with Cronbach's alpha and an item analysis. Findings also indicate that the scale scores from 6 subscales are considered normal distributed (e.g. presence) whereas the scale scores from 3 subscales are considered negatively skewed (e.g. skill). This study provides important new insight into UX in IVEs assessment.

Index Terms - **Virtual Reality; Immersive Virtual Environment; User Experience; UX Questionnaire; Evaluation.**

I. INTRODUCTION

After over 10 years of absence in the media landscape, Virtual Reality (VR) interest resumed in early 2012. Since, there is a rise in VR research to face the increase of new technology emergence. Nevertheless, according to us, the actual User eXperience (UX) models for Virtual Environments (VE) discussed in scientific literature do not include the whole UX in VE key components.

This led us to propose a definition of the UX in Immersive Virtual Environments (IVE) that takes into account the multiple facets of the UX in several fields of VR (entertainment, education, edutainment). Along with this new definition, we designed a new UX in IVE holistic model (Figure 1) and de-

signed a suitable measurement method based on our model (Tcha-Tokey et al., 2015). The model is based on the key components recommended by the literature and the designed measurement method is a questionnaire based on the components of our model. Our questionnaire is designed on the basis of existing questionnaires, since most of the UX in IVE components of our model can be measured through general (non-VR specific) UX questionnaires or specific UX for VR questionnaires.

This paper describes the construction and the validation of our UX questionnaire for IVE (IVEQ). Indeed, we want to verify that the items selected to design our questionnaire measure properly the UX components from the original questionnaires. Firstly, we review the selected questionnaires we based our own questionnaire on. Secondly, we describe the experiment that took place to validate our questionnaire and finally we discuss the reliability and the sensitivity of our questionnaire.
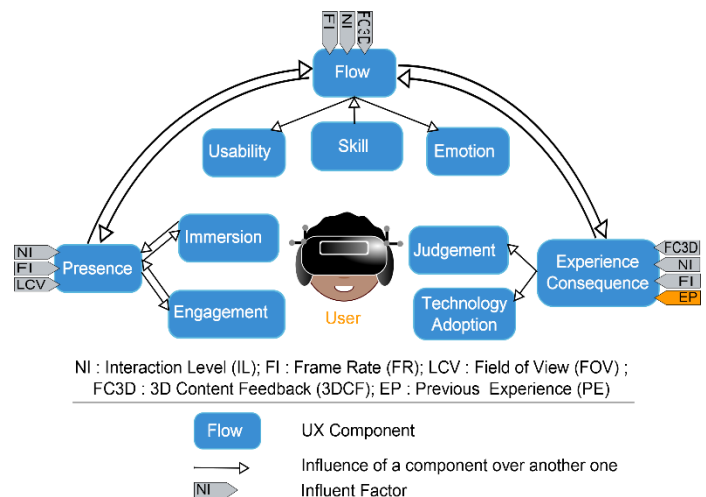

Figure 1: Our holistic User Experience in Immersive Virtual Environment model.

E-mail: katy.tcha-tokey@ensam.eu

II. PROPOSITION OF A UNIFIED UX QUESTIONNAIRE

The UX is defined by a variety of different components depending on the field the experience is lived in. In our study, we define the UX through 10 components. These components also structure our UX questionnaire in 10 subscales (Appendix 1). Our questionnaire was designed on the basis of 9 UX questionnaires. We define the 10 subscales that compose our questionnaire in section 2.1. Two steps led us to our unified questionnaire: the UX questionnaires selection and the items selection. The questionnaires selection criteria are detailed in section 2.1. The items selection criteria and the internal structure of our questionnaire are detailed in section 2.2.

2.1 User Experience Questionnaires Review

The UX in IVE can be measured by either subjective methods or objective methods. Yet a combination of both methods might provide results that are more reliable (Yao et al., 2014). Subjective methods (e.g. questionnaire, interviews, focus groups...) provide results through the user's point of view, attitudes or preferences, whereas objective methods (e.g. electroencephalogram, electromyogram, time completion, level reached...) provide results through observable evidence. For a complete diagnosis of UX with the user's thoughts patterns and believes, it is preferable to combine the objective data with subjective data using suitable tools such as a questionnaire and "UX heat maps" (Georges et al., 2016). However, objective methods used today to assess user experience are questionable due to signal contamination: motor interference in Brain Computer Interaction (BCI) and Electromyogram (EMG) (McFarland et al., 2005), signal distinction of two close emotions such as stress or excitement in Electrocardiogram (ECG) or Electroencephalogram (EEG) (Chai et al., 2014; Kim and Chang, 2015), late response latency in Galvanic Skin Response (GSR) and Skin Temperature (SKT) (Kim and Chang, 2015).

Considering the limits of objectives methods and the advantages of questionnaires, UX specialists tend to opt for the use of questionnaires. Indeed, literature shows that the questionnaire is, currently, the most commonly used method for the measure of UX components (e.g. presence, engagement, immersion, flow, emotion, judgement...), and a large number of questionnaires have been proven valid and reliable. For all these reasons, we focus our approach on the questionnaire method in this paper. Actually, to the best of our knowledge, no UX questionnaire integrates all of the UX key components for IVE concern.

In order to help us choose the suitable questionnaires we defined two questionnaires selection criteria:

- Validity of the questionnaire (i.e. the whole validation process of the questionnaire is published in a paper).

- Frequent use of the questionnaire or the questionnaire is based or inspired by a frequent used questionnaire (i.e. we consider "frequent" the use of a questionnaire which is cited at least by 20 other papers in the scientific literature).

After defining the criteria, we found a suitable questionnaire for each of the 10 UX components we proposed (presence, engagement, immersion, flow, usability, skill, emotion, experience consequence, judgement and technology adoption) to measure the overall UX in IVE (Tcha-Tokey et al., 2016). The suitable questionnaires and the UX components are defined below:

- Presence is a component defined as the user's "sense of being there" in the VE. The concept of presence can be divided into two categories: physical presence in the virtual environment and social presence in the collective or collaborative virtual environment (Pallot et al., 2013). Most measures of presence try to address both. Engagement is a component defined as the "energy in action, between a person and its activity consisting of a behavioral, emotional and cognitive form". The Presence Questionnaire (PQ) created by Witmer and Singer measures presence and engagement (Witmer and Singer, 1998), it identifies the degree to which individuals experience presence and engagement in VE. This questionnaire is composed of 24 items divided in 5 subscales: involved, natural, auditory, resolution and interface quality.

- Immersion is a component defined as the "illusion" that "the virtual environment technology replaces the user's sensory stimuli by the virtual sensory stimuli". The Immersion Tendency Questionnaire (ITQ) created by Witmer and Singer measures immersion (Witmer and Singer, 1998), it identifies the tendency of individuals to be immersed. This questionnaire is composed of 16 items divided in 3 subscales: involvement, focus and game.

- Flow is a component defined as "a pleasant psychological state of sense of control, fun and joy" that the user feels when interacting with the VE. The Flow4D16 questionnaire created by Heutte measures the flow component (Heutte and Fenouillet, 2015). It identifies the degree with which the user is absorbed by his task. The questionnaire consists of 16 items divided in 4 subscales: cognitive absorption, altered time perception, lack of self-preoccupation, well-being.

- Skill is a component defined as the knowledge the user gain in mastering his activity in the virtual environment. The Computer Self-Efficacy (CSE) questionnaire created by Murphy measures the skill component (Murphy, Coover, and Owen, 1989). It identifies the attitude of a user toward a computer technology, the degree with which he feels comfortable with a computer. This questionnaire is a reference in the education field to evaluate adult student's computer skills. The questionnaire consists of 32 items with 3 subscales representing different levels assessment of computer skills: beginning, advanced, mainframe.

- Emotion is a component defined as the feelings (of joy, pleasure, satisfaction, frustration, disappointment, anxiety ...) of the user in the VE. The Achievement Emotions Questionnaire (AEQ) created by Pekrun measures the emotion component (Pekrun et al., 2011). It identifies the emotion experienced in achievement situations.

There are 3 subscales representing 3 situations: class-related, learning-related and test-related. This questionnaire is based on 9 emotions: enjoyment, hope, pride, relief, anger, anxiety, shame, hopelessness, boredom. The questionnaire consists of 232 items. It proposes a large number of situations that matches or that can be easily translated in a situation such as a user being in a VE.

- Usability is a component defined as the ease of learning (learnability and memorizing) and the ease of using (efficiency, effectiveness and satisfaction) the VE. The System Usability Scale (SUS) created by Brooke measures the usability component (Brooke, 1996). This scale has been created on a base of 50 usability questionnaires. It identifies "the appropriateness of a purpose", in other words, it identifies if the way we propose to use our VE is appropriate. The questionnaire consists of 10 items and is unidimensional.

- Technology adoption is a component defined as the actions and decisions taken by the user for a future use or intention to use of the VE. The Unified Technology Acceptance and Use of Technology (UTAUT) questionnaire created by Venkatesh and al. measures the technology adoption component (Venkatesh et al., 2003). It identifies the degree with which the user will adopt and use the system, in other words, the likelihood of success for new technology introduction. This questionnaire consists of 31 items divided in 8 subscales: performance expectancy, effort expectancy, social influence, facilitating conditions, attitude toward using technology, self-efficacy, anxiety, behavioral intention to use the system.

- Judgement is a component defined as the overall judgement of the experience in the VE. The AttracDiff questionnaire created by Hassenzahl, Burmester and Koller measures the judgement component (Hassenzahl, Burmester, and Koller, 2003). It identifies the user's attraction in a pragmatic and hedonic way towards the system. This questionnaire consists of 28 items divided in 4 subscales: perceived pragmatic quality, perceived hedonic quality-stimulation, perceived hedonic quality-identification, attractiveness.

- Experience Consequence is a component defined as the symptoms (e.g. the simulator sickness, stress, dizziness, headache …) the user can experience in the VE. The Simulator Sickness Questionnaire (SSQ) created by Kennedy measures the experience consequence component (Kennedy et al., 1993). It identifies the negative consequences the user can have while using the IVE. These negative consequences are assessed through 16 items divided in 3 subscales: nausea, oculomotor problems, disorientation.

2.2 A Unified UX Questionnaire

2.2.1 Structuring of the Instrument

The questionnaire we designed aims at measuring the various facets of the UX in IVE. The idea is to offer to the user, after a certain amount of time in the VE, one unique questionnaire that measures the whole 10 components of the UX. So, the questionnaire is composed of a set of items that gathers the user's opinion, beliefs and preferences on the VE that he/she experienced in terms of presence, engagement, immersion, flow, usability, skill, emotion, experience consequence, judgement and technology adoption. These items provided from the nine following questionnaires: PQ, ITQ, Flow4D16, CSE, AEQ, SUS, UTAUT, AttrakDiff, SSQ.

Our questionnaire is composed of 87 items. Among them, 75 items have a 10-point Likert scale, 12 items (grouped in four) have a differential scale items and 3 items are open questions. To select these 87 items, we relied on a single criterion: the meaning of the chosen items had to be different enough from each other (even if they measure the same component), so that the user does not find the items redundant.

This questionnaire comprised ten subscales as described below. Presence was assessed using 12 items (e.g. "*The virtual environment was responsive to actions that I initiated*") adapted from PQ scales (Witmer and Singer, 1998). Engagement was assessed using 3 items (e.g. "*The sense of moving around inside the virtual environment was compelling*") adapted from PQ scales (Witmer and Singer, 1998). Immersion was assessed using 7 items (e.g. "*I felt stimulated by the virtual environment*") adapted from ITQ scales (Witmer and Singer, 1998). Flow was assessed using 11 items (e.g. "*I felt I could perfectly control my actions*") adapted from Flow4D16 scales (Heutte and Fenouillet, 2015). Usability was assessed using 3 items (e.g. "*I thought the interaction devices (oculus headset, gamepad and/or keyboard) was easy to use*") adapted from SUS scales (Brooke, 1996). Emotion was assessed using 14 items (e.g. "*I enjoyed being in this virtual environment*") adapted from AEQ scales (Pekrun et al., 2011). Skill was assessed using 6 items (e.g. "*I felt confident selecting objects in the virtual environment*") adapted from CSE scales (Murphy, Coover, and Owen, 1989). Judgement was assessed using 12 items (e.g. "*Personally, I would say the virtual environment is impractical/practical*") adapted from AttracDiff scales (Hassenzahl, Burmester, and Koller, 2003). Experience consequence was assessed using 9 items (e.g. "*I suffered from fatigue during my interaction with the virtual environment*") adapted from SSQ scales (Kennedy et al., 1993). Technology adoption was assessed using 9 items (e.g. "*If I use again the same virtual environment, my interaction with the environment would be clear and understandable for me*") adapted from UTAUT Scales (Venkatesh et al., 2003). We added 3 open questions at the end of the questionnaire to allow the user to express the positive as well as the negative experience he wish to share and the improvements he wish to provide to the environment.

We made some arrangement to existing items in order to create a questionnaire better related to VE:

- In the PQ, some subscales (e.g. IFQUAL: Interface Quality; NATRL: Natural, AUD: Auditory, RESOL: Resolution) only had 2 or 3 items, in that case we did not have to make a selection and picked all of the items of the subscale (e.g. AUD:"*14 - I correctly identified sounds produced by the virtual environment.*"; "*15 - I correctly localized sounds produced by the virtual environment.*").

- In the ITQ, the items in the subscale GAMES could hardly apply to VE and the items from the involvement (INVOL) and FOCUS subscales did not apply right away to our context. Therefore, we chose items that could easily be adjusted to our context, and rewrote the items when necessary (e.g. "*How mentally alert do you feel at the present time?*" becomes "*16 - I felt mentally alert in the virtual environment.*", "*How frequently do you find yourself closely identifying with the characters in a story line?*" becomes "*18 - I identified to the character I played in the virtual environment.*" . . .).

- In the AEQ, one subscale could hardly apply to VE (i.e. Test emotion). For the 2 remaining subscales we chose to select 3 items in 3 emotion categories (positive activating: enjoyment, negative activating: anxiety, negative deactivating: boredom).

- In the UTAUT questionnaire, 5 direct determinants of intention subscales could hardly apply to VE or were redundant with other items already selected (i.e. performance expectancy, social influence, self-efficacy, anxiety, behavioral intention to use the system). Three of the subscales did apply to VE (i.e. effort expectancy, attitude toward using technology, facilitating conditions).

- We adjust most of the items we selected so that they could fit perfectly to VE. In some cases, changing the words "system" or "class" to "virtual environment" was enough (e.g. "*I enjoy being in class*" becomes "*37: I enjoyed being in the virtual environment*"), in other cases we did adjust the whole item to apply for VE (e.g. "*I feel confident making selections from an onscreen menu.*" becomes "*52 - I felt confident selecting objects in the virtual environment.*")

2.2.2 Answer Modalities and Scoring

The participants' UX scores were collected through a 10-point Likert scale (1 = strongly disagree, 10 = strongly agree) for 75 items. For the 12 items (grouped in 4) of the judgement, scores were collected through a semantic differential scale: point 1 was coded as a negative-connoted adjective (e.g. impractical, confusing, amateurish . . .) whereas point 10 was coded as a positive-connoted adjective (e.g. practical, clear, professional . . .).
A high score in a subscale means that the UX component measured is highly perceived by the participant (e.g. Presence equals to 9 means that the participant felt really present, "he felt there", while he was in the Virtual Environment). A low score in a subscale means that the UX component measured is poorly perceived by the participant (e.g. Presence equals to 2 means that the participant did not really felt present, while he was in the Virtual Environment, there was few or no moments where he would easily forget about the real environment).

2.2.3 Execution Time

This questionnaire execution time varies from 15 to 20 minutes.

III. Validation of a unified UX questionnaire

To validate our unified UX questionnaire described in the previous section in terms of reliability and sensitivity, we conducted a study with an edutainment IVE prototype called "Think and Shoot" with 116 participants.

3.1 Aim of our Study

Our study aims at providing a UX questionnaire available for IVEs designers to help them in the assessment and the improvement of the UX in IVEs (prototypes or final solutions).

3.2 Method

3.2.1 Participants

116 participants (25 women and 91 men) aged 18-63 years (M = 24.6, SD = 7.55) took part in the study. Among them, 9 participants did not answer the activity question and the 107 others were distributed as follows:

- 88 participants work or study in Information and Communications Technology (ICT) or Computer Science fields (e.g. VR engineers, VR research engineers, network administrator, web developer, web designer, graphic designer, PhD student in VR, master degree undergraduates in VR, professional degree undergraduates in ICT, technical degree undergraduates in Multimedia and Internet . . .).

- 19 participants work or study in various other fields (e.g., specialized education, marketing, food service, military, public relations, Bank . . .).

Concerning their experience level with 3D technologies, 34 participants are considered experienced with 3D technologies, they scored at least 21/42 points (see section 3.2.4 for scoring details) in the 3D technology expertise survey (M = 26.79, SD = 5.09), they use 3D technologies in a regular basis to play 3D video games or to create 3D contents. 82 participants are considered non-experienced, they scored less than 21/42 points in the 3D technology expertise survey (M = 10.97, SD = 5.11), they never use or use few 3D technologies dedicated to create 3D contents or they never (or very little) play with 3D video games (Table 1).

Profile	Level of experience	Participants
Experienced	> or = 21 points in the 3D technology expertise survey	34
Non-Experienced	<21 points in the 3D technology expertise survey	82

Table 1: Number of experienced versus non-experienced participants in 3D Technologies.

3.2.2 Procedure

The experiment took place in a 16 m^2 room (Figure 2) which was rearranged for the experiment purpose during two months (from the 1st of february of 2016 to the 31st of march 2016).
The experiment had three steps:

- During the first step of the experiment, we installed the participants in the experiment room and asked them to read and sign a consent document presenting the laboratory and the experiment confidentiality rules. We then asked them to complete a "participant identification survey".

- During the second step of the experiment, first, we explained the whole experiment goal to the participants. Secondly, we explained the training goal and we asked the participants to put on the Oculus and the audio headsets for a training session of about 5 minutes (the participants could ask for more or less training time if they felt more or less comfortable in the IVE).

- During the third step of the experiment, we explained the regular session goal to the participants. We then asked them to put on the Oculus and audio headsets for the regular session of 5 minutes. After the end of the session, the participants completed our UX questionnaire. Each participant spent between 30 to 45 minutes in the experiment room.

Figure 2: The 16m^2 room dedicated to the experiment purpose.

3.2.3 Material and Measures

Three questionnaires have been used: a consent document, an identification questionnaire and our UX questionnaire.

A consent document was used to inform the participant about the laboratory activity and to collect his agreement to participate in our experiment under the announced conditions (e.g. recorded experiment, confidentiality...). This document required the participant's personal information (i.e. name, date of birth, address, occupation). A participant identification survey was used to collect the user's skills. The participant's last diploma, and current diploma or occupation were asked. Three items with a 5-point Likert scale were dedicated to programming expertise (0 = No knowledge, 4 = Excellent knowledge). Two multiple-choice questions were used to assess the participant's ability to recognize a function and a parameter in an instruction: one point was given for a correct answer and zero point was given for a wrong answer. Two matrix scale questions were dedicated to technology expertise (0 = Never, 1 = Little, 2 = Sometimes, 3 = Often, 4 = Very often). The first matrix scale question was dedicated to the usage frequency of interaction devices such as VR headset, gamepad, joystick, Kinect, leap motion... and the second matrix scale question was dedicated to usage frequency of 3D video games and 3D software such as 3D scene design software (i.e. Virtools, Unity...), modelling software (i.e. 3DSmax, Maya ...), CAO software (i.e. AutoCAD, Architectural desktop ...).

Our UX in IVE questionnaire of 87 items and 3 open questions is used to assess the UX. All items and questions were originally in French. As described previously, our UX questionnaire consists of 12 items to measure presence, 3 items to measure engagement, 7 items to measure immersion, 11 items to measure flow, 3 items to measure usability, 6 items to measure skill, 15 items to measure emotion, 9 items to measure experience consequence, 12 items (grouped in 4) to measure judgement and 9 items to measure technology adoption.

The study was conducted with the edutainment IVE prototype "Think and Shoot" aiming at collecting three types of balls and to shoot on two types of evil creatures pursuing the participants, according to the instructions given on a panel displayed inside the VE. (Figure 3).

"Think and Shoot" is designed with the development tool UNITY© and displayed in an Oculus development kit 2 (DK2). To interact with the VE, a Logitech wireless gamepad is used. The 3D spatialized sound is rendered in a Tritton AX 180 audio headset. The VE is launched on a Dell 64bits with 4GB of RAM computer and an Intel® Xeon® processor CPU E5-16030 2.80GHz. The computer operating system is Windows 10 Professional. Oculus runtime SDK 0.7 and NVIDIA 356.04 GeForce win10 drivers were installed.

The VE factors were fixed to a field of view of 106 degrees and a framerate of 70 FPS –recommended by the Oculus Best Practices (Oculus©, 2016). The Gamepad allowed the user to move forward, backward and on the sides with one of the joysticks, and with the other, he could rotate on himself. The user had a minimap of the environment to help him localize the balls.

Figure 3: Virtual edutainment environment screen shot of a fire evil creature in the first level.

3.2.4 Task

We proposed a training session and a regular session to the participants. In the training session, after collecting the balls, the participant earns a point if he/she shoots correctly a fire ball on the blue sphere target, an ice ball on the red sphere target and a lightning ball on the green sphere target. During the regular session, after collecting the balls, the participant earns a point if he/she shoots correctly an ice ball on the fire evil creature, a fire ball on the ice evil creature. If he/she shoots a lightning ball on both evil creatures, they are then frozen and they cannot move forward anymore. There are six levels in the edutainment VE.

During the first level, the participant can only shoot ice balls and only one fire evil creature pursues him/her. In the second level, there are more fire evil creatures. In the third level, the participant can shoot both ice balls and lightning balls. In the fourth and fifth levels, the participant can shoot the three types of balls (ice, fire, and lightning) and he/she can shoot on both ice and fire evil creatures.

3.3 Data Analysis

3.3.1 Psychometric Properties

Questionnaires are self-appraisal methods and thus induce two kinds of problems: the misunderstanding of the items' meaning and the risk of giving a stereotypical answer. So, it is necessary to analyze their items and subscales quality through three recommended psychometric properties (Fernandez et al., 2005): the validity, the reliability and the sensitivity.

The validity refers to the accuracy of the affirmations that can be done through test scores. There are various types of validity, for example the construct validity assessed by analyzing the factor structure or the concurrent validity assessed by correlating a test with other similar tests. The reliability assesses the consistency of a measure. A highly reliable measure is a measure that produces the same result under consistent conditions. There are various types of reliability, for example the test-retest assessed by observing stability of the results throughout time or the internal consistency assessed by measuring the Cronbach's alpha or the item correlation (through an item analysis). The sensitivity measures the ability of an evaluation method to detect different enough results among individuals. The sensitivity is assessed by observing the scores distribution according to the normal distribution. The sensitivity can be intraindividual or interindividual.

According to the sample size, we meet the requirements to assess the reliability and the sensitivity of our UX questionnaire. For the validity, we observed that the sample size of our experiment (N=116) did not match with the sample size required in the literature – N=200 (Anderson and Gerbing, 1987), N = 300 (Tabachnik and Fidell, 2007), 10 or more participants per item (Nunnally, 1978). Indeed, according to various studies on factor analysis (Comrey and Lee, 1992; MacCallum et al., 1999) there is an effect of sample size on factor analysis: "*As N increases, sampling error will be reduced and sample factor analysis solutions will be more stable and will more accurately recover the true population structure*". This does not make possible the validation of our questionnaire with a factor analysis study.

3.3.2 Reliability

In order to determine the reliability, we calculated the internal consistency of each subscale (e.g., presence, engagement, immersion, flow . . .) with Cronbach's alpha and we calculated the Pearson product-moment correlation coefficient (PCC) of each item through an item analysis. Regarding the Cronbach's alpha a value of 0.70 is recommended to consider a measure as being reliable – as a standard norm (Devellis, 2003). We then chose the item analysis to continue with the reliability analysis and check if each item score is consistent compared to the global score of our UX questionnaire. We used Cohen's convention (Cohen, 1988) to interpret the values: a correlation coefficient of 0 indicates that there is no relation; a correlation coefficient of 1 indicates a perfect positive correlation; a correlation coefficient of -2 indicates a perfect negative correlation; a correlation coefficient between 0.1 and 0.29 indicates a weak correlation; a correlation coefficient between 0.3 and 0.49 indicates a moderate correlation and a correlation coefficient between 0.5 and 1.0 indicates a strong correlation.

3.3.3 Sensitivity

In order to determine the sensitivity, we calculated the interindividual sensitivity; that is to say, the scores distribution, according to the normal distribution, among different individuals using qualitative observation and the Kolmogorov-Smirnov test (K-S test). The distribution can vary from symmetric distribution (no skew), positively skewed distribution or negatively skewed distribution.

IV. RESULTS

We present below the results of the internal consistency with Cronbach's alpha and an item analysis to check the reliability of our UX questionnaire. In addition, we present the result of the scores distribution on distribution graphs and with an analysis of the K-S test to check the interindividual sensitivity of our UX questionnaire. These psychometric properties are calculated with the IBM® SPSS® Statistics software.

4.1 Reliability

4.1.1 Internal Consistency: Subscale Reliability

Questionnaire reliability data (Cronbach's alpha α) of our UX questionnaire are presented in Table 2. These data indicate for presence, engagement, immersion, flow, skill, emotion, experience consequence, judgement and technology adoption subscales satisfactory internal consistency ($\alpha > 0.7$; i.e. α from 0.718 to 0.908). Whereas for usability subscale the data indicate unsatisfactory internal consistency ($\alpha < 0.7$; i.e. $\alpha = 0.465$).

Table 2 presents main results concerning internal consistency. Four items were dropped because of their low Cronbach's alpha.

Results (N=116)

Subscales	Cronbach's α	No. of items	Sources	Dropped items
Presence	0.755	11	Witmer and Singer 1998	11
Engagement	0.759	3	Witmer and Singer 1998	.
Immersion	0.767	7	Witmer and Singer 1998	.
Flow	0.826	11	Heutte 2010	.
Usability	0.465	3	Brooke 1996	34, 35, 36
Skill	0.820	6	Murphy 1989	.
Emotion	0.718	14	Pekrun 2011	.
Experience consequence	0.908	9	Kennedy 1993	.
Judgement	0.804	12	Hassenzahl, Burmester and Koller 2003	.
Technology adoption	0.781	9	Venkatesh 2003	.

Table 2: Results of Cronbach's alpha for our questionnaire's sub-scales.

4.1.2 Internal Consistency: Item Analysis

Item analysis data of our UX questionnaire are presented in Table 3. The data indicate satisfactory Pearson product-moment Correlation Coefficient (PCC) for 68 items out of 87 meaning that these items are significantly correlated (moderately to strongly correlated) with the global score of our UX questionnaire, and an unsatisfactory PCC for 15 items out of 87 meaning that these items are weakly correlated with the global score of our UX questionnaire.

More precisely, Table 3 shows that 9 out of 11 items from the presence subscale, each of the 3 items from the engagement subscale, 5 out of 7 items from the immersion subscale, 10 out of 11 items from the flow subscale, 11 out of 15 items from the emotion subscale, each of the 6 items from the skill subscale, 9 out of 12 items from the judgement subscale, 8 out of 9 items from the experience consequence subscale, 7 out of 9 items from the technology adoption subscale are significantly correlated. Fifteen items were dropped because of their low PCC.

Items	PCC
Item1	0.233*
Item2	0.491**
Item3	0.482**
Item4	0.369**
Item5	0.535***
Item6	0.491**
Item7	0.317**
Item8	0.344**
Item9	0.520***
Item10	0.516***
Item12	0.149*
Item13	0.302**
Item14	0.342**
Item15	0.301**
Item16	0.593***
Item17	0.390***
Item18	0.226*
Item19	0.387**
Item20	0.640**

Table 3: Pearson Correlation Coefficient (PCC) for items 1-20 of our UX questionnaire.
Note. *Weak correlation; ** Moderate correlation; ***Strong correlation; item 11 was dropped for a better α in the presence subscale.

Items	PCC
Item21	0.187*
Item22	0.501***
Item23	0.493**
Item24	0.356**
Item25	0.527***
Item26	0.462***
Item27	0.593***
Item28	0.475**
Item29	0.286*
Item30	0.360**
Item31	0.653***
Item32	0.724***
Item33	0.454**
Item37	0.737***
Item38	0.111*
Item39	0.668***
Item40	0.650***
Item41	0.055***
Item42	0.116*
Item43	0.249*
Item44	0.084**
Item45	0.437**
Item46	0.472**
Item47	0.458**
Item48	0.529***
Item49	0.132**
Item50	0.163*
Item51	0.585***
Item52	0.487**
Item53	0.485**
Item54	0.354**
Item55	0.360**
Item56	0.489**
Item57	0.417**
Item58 A	0.673***
Item58 B	0.603***
Item58 C	0.550***
Item59 A	0.281*
Item59 B	0.453**
Item59 C	0.289*
Item60 A	0.367**
Item60 B	0.488**
Item60 C	0.426**
Item61 A	0.461**
Item61 B	0.515***
Item61 C	0.290*
Item62	0.444**
Item63	0.504***
Item64	0.426**
Item65	0.322**
Item66	0.281*
Item67	0.471**
Item68	0.490**
Item69	0.492**
Item70	0.445**
Item71	0.573***

Table 3 Continued: Pearson Correlation Coefficient (PCC) for items 21-71 of our UX questionnaire.
Note. *Weak correlation; ** Moderate correlation; ***Strong correlation; items 34, 35, 36 were dropped due to unsatisfactory reliability of the usability subscale.

Items	PCC
Item72	0.595***
Item73	0.567***
Item74	0.453**
Item75	0.429**
Item76	0.459**
Item77	0.424**
Item78	0.280*
Item79	0.236*

Table 3 Continued: Pearson Correlation Coefficient (PCC) for items 72-79 of our UX questionnaire.
Note. *Weak correlation; ** Moderate correlation; ***Strong correlation.

4.2 Sensitivity

Questionnaire sensitivity data of our UX questionnaire are presented in Figures 4-12. The data indicate for the presence (Figure 5), engagement (Figure 7), immersion (Figure 6), flow (Figure 4), emotion (Figure 9) and judgement (Figure 10) subscales approximatively symmetrically distributed (no skew) scores according the normal distribution. Whereas the scores of the skill (Figure 8), technology adoption (Figure 11) and experience consequence (Figure 12) subscales are negatively skewed.
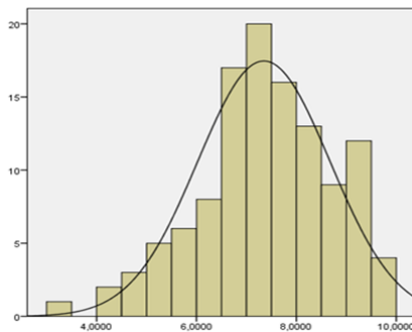

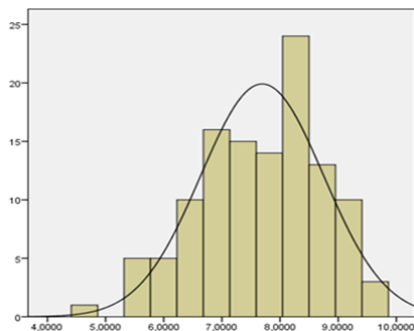Figure 4: Flow scores distribution.

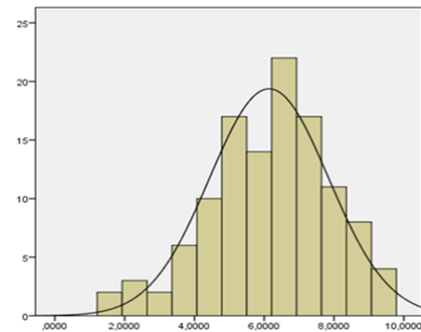Figure 5: Presence scores distribution.

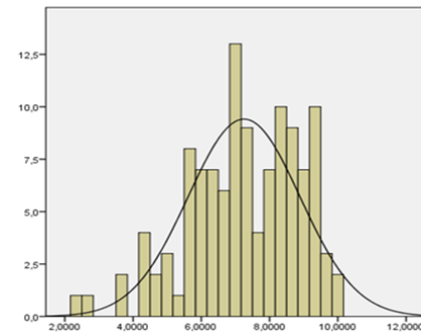Figure 6: Immersion scores distribution.
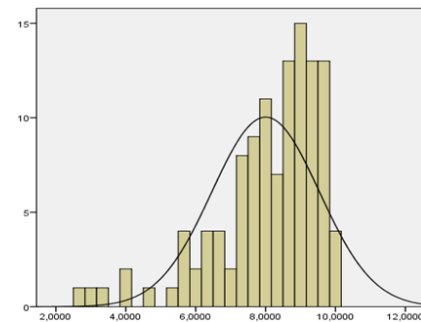

Figure 7: Engagement scores distribution.

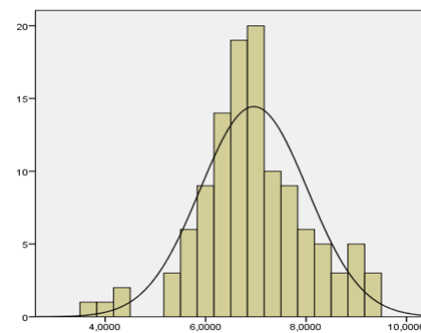Figure 8: Skill scores distribution.

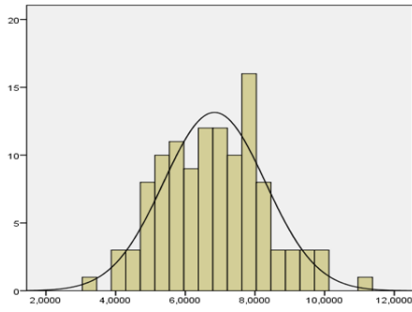Figure 9: Emotion scores distribution.
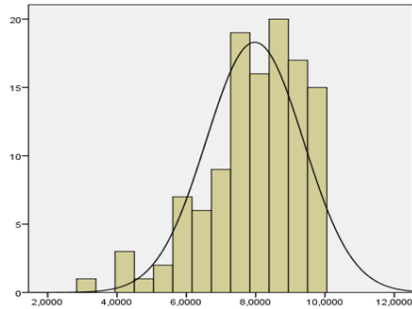
Figure 10: Judgement scores distribution.

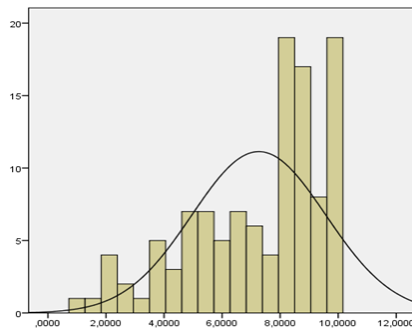Figure 11: Technology adoption scores distribution.

Figure 12: Experience consequence scores distribution.

The asymmetrical observation of the skill, technology adoption and experience consequence subscales is confirmed by the asymmetry values in Table 4.

The asymmetry might be explained by the high number of young individuals (M = 24.6 years) skilled in video games (70 individuals out of 116 often or very often play to 3D video games). Actually, the majority of the participants tend to be skilled in the VE (explained by a majority of positive scores in the skill subscale). They are more likely to adopt 3D technologies as well (explained by a majority of positive scores in the technology adoption subscale) and they tend to be less sick inside the VE (explained by a majority of positive scores in the experience consequence subscale).

Results (N=116)						
Subscales	**M**	**SD**	**Variance**	**Asymetry**	**K-S Z**	**p**
Presence	7.692790	1.0565509	1.116	-0.353	0.839	0.483
Engagement	7.252874	1.6386755	2.685	-0.570	1.062	0.209
Immersion	6.139163	1.7060320	2.911	-0.368	0.740	0.644
Flow	7.349373	1.3249692	1.756	-0.472	0.608	0.853
Skill	8.001437	1.5372187	2.363	-1.390	1.393	0.041
Emotion	6.959770	1.0680744	1.141	-0.062	1.084	0.190
Experience consequence	7.265326	2.3089256	5.331	-0.822	1.809	0.003
Judgement	6.838362	1.4659290	2.149	0.208	0.491	0.969
Technology adoption	7.975096	1.4040878	1.971	-0.925	1.014	0.256

Table 4: Mean (M), Standard Deviation (SD), Variance, Asymmetry Kolmogorov-Smirnov Z (K-S Z) and p of the scores of our questionnaire subscales.

4.3 Validated Questionnaire

These data allow us to successfully end with a new version of our validated UX questionnaire according to reliability and sensibility. We dropped 19 items from the original 87 items of our questionnaire. The IVEQ is now composed of 68 items divided in 9 subscales and 3 open questions. 9 items compose the presence subscale; 3 items compose the engagement subscale; 5 items compose the immersion subscale; 10 items compose the flow subscale; 11 items compose the emotion subscale; 6 items compose the skill subscale; 9 items compose the judgement subscale; 8 items compose the experience consequence subscale; 7 items compose the technology adoption subscale. See Appendix 1 for the full version of the IVEQ.

V. DISCUSSION

This present research set out to integrate the fragmented theory and research on UX into a unified UX in IVE questionnaire (IVEQ). Our study enabled us to assess the reliability of 9 subscales of our UX questionnaire specifically through the internal consistency, provided the rejection of the usability subscale (result show unreliability in the usability subscale with $\alpha = 0.465$ and satisfactory reliability for the other 9 subscales with $\alpha = 0.718 - 0.908$). With the internal consistency analysis, 4 items were dropped because of their low Cronbach's alpha (1 item from presence and 3 items from usability). The usability unsatisfactory result can be explained in several ways: a small amount of items chosen to measure the usability, the items were not enough

adjusted to the context or the items chosen were too redundant with other items. With the item analysis, 15 items were dropped because of their low correlation coefficient suggesting that they contributed relatively little to the reliability of our questionnaire (2 other items from presence, 2 items from immersion, 1 item from flow, 4 items from emotion, 3 items from judgement, 1 item from experience consequence, 2 items from technology adoption).

Earlier studies have provided evidence of the internal consistency of the original questionnaires we used to create our own. In comparison in Table 5, 7 subscales of our questionnaire have a slightly lower α, 2 subscales have a slightly higher α and one subscale have a strongly lower α. Our findings concerning the reliability of our questionnaire are in certain aspects similar to the original studies (Table 5).

Subscales	Original questionnaires	α in our study	α in original questionnaires	Sources
Presence	PQ	0.755	0.88	Kennedy et al. 1993
Engagement	PQ	0.759		
Immersion	ITQ	0.767	0.81	
Flow	Flow4D16	0.826	0.84 - 0.86	Heutte et al. 2010
Skill	CSE	0.820	0.95	Murphy et al. 1989
Usability	SUS	0.465	0.92	Lewis and Sauro 2009
Emotion	AEQ	0.718	0.78 - 0.93	Pekrun et al. 2011
Experience consequence	SSQ	0.908	0.71	Bailenson and Yee 2006
Judgement	AttracDiff	0.804	0.73 - 0.90	Hassenzahl et al. 2003
Technology adoption	UTAUT	0.781	0.87 - 0.91	Venkatesh et al. 2003

Table 5: Comparison of Cronbach's alphas in our UX questionnaire and in the original questionnaires.

So far, only a limited number of studies investigated the sensitivity of a UX questionnaire. We assessed the sensitivity through observation and the K-S test and found scores normally distributed for 6 subscales (i.e. presence, engagement, immersion, flow, emotion, judgement), and negative skewed distribution for 3 subscales (i.e. skill, technology adoption and experience consequence). These 3 subscales negatively skewed can be explained by a high number of individuals skilled in video games that tend to be skilled in the VE as well, that are likely to adopt 3D technologies more easily and that tend to be less sick inside the VE.

Even if our questionnaire shows a good reliability and a good sensitivity according previous criteria, the validation of our UX questionnaire could be improved through further studies.

The first limitation concerns the lack of investigation of other reliability parameters such as the test-retest method, due to the already big workload for the participants requested by our experimental protocol. Indeed, the load requested by the experiment might have caused more risks of errors and fatigue in the experiment process and more risk of random answers in the questionnaire. Nevertheless, measuring the reliability along time is feasible now that our questionnaire does not need so much validation. We might consider an experiment with participants during two days, the exact same experiment with the exact same participants will take place on the first and the second day.

A second limitation concerns the lack of investigation of other sensitivity parameters such as the intraindividual sensitivity for the same reason as the test-retest method.

A third limitation concerns the lack of investigation of the validity parameters such as the construct validity due to the unsatisfactory sample size required or the criterion validity due to the incompatibility with our experimental protocol. Indeed, the criterion validity method requests a comparison between our unified questionnaire and the original questionnaires. Unfortunately, no dataset of the original questionnaires was collected because the experiment would have been very cumbersome for the participants (i.e. the experiment would have requested a completion of 10 questionnaires for each participant).

VI. CONCLUSION

This paper provides an instrument to assess the multiple facets of UX in IVEs. It should be noted that this paper's objective is to present and validate a subjective method: our unified UX questionnaire, the IVEQ. Once validated, our questionnaire's data might be combined at will with other objective data. According to previous criteria, our unified UX questionnaire is a reliable and sensitive instrument. For the reliability, we were able to validate the internal consistency with Cronbach's alpha (except for one subscale) and item correlation that reinforces the reliability of our UX questionnaire subscales (provided 19 items dropped). For the sensitivity, we observed the normal distribution of the scores for 6 subscales and the negative skewed distribution of the scores for 3 subscales in our UX questionnaire, due to the high number of skilled participants with VE.

Some questions remain on the unsatisfactory internal consistency result of the usability subscale, and on the validity analysis of our UX questionnaire. Indeed, a higher number of participants would have helped enrich and refine the whole validation process (i.e. construct validity, concurrent validity, confirmatory factor analysis) even if we consider our validation process suitable for experiments with less than 200 participants.

This study provides important new insight into UX in IVEs assessment. Our unified questionnaire is general in terms of subscales and items. This suggests two consequences: it may be used for different types of VE such as therapeutic, design or collaborative applications (even if we used it for an edutainment VE in our study), and for different types of devices such as a Cave Automatic Virtual Environment (CAVE), z-Space, . . . (even if we used a HMD in our study). Finally, the present study has a number of important implications for UX design process. First, our UX in IVE questionnaire can be used in the earlier phases of a VE design. The questionnaire might be used as soon as a prototype is designed to assess the UX and therefore favor a better UX for the final solution. Secondly, designers can use the UX in IVE questionnaire as a guide, to help them focus on the unsatisfactory aspects of the VE in terms of UX, knowing that way, which aspect of the UX has to be improved to provide a greater and suitable experience for customers.

ACKNOWLEDGEMENTS

REFERENCES

Anderson, J. C., and Gerbing, D. W. 1987. The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika* 49(2):155–173.

Bailenson, J. N., and Yee, N. 2006. A longitudinal study of task performance, head movements, subjective report, simulator sickness, and transformed social interaction in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments* 15(6):699–716.

Brooke, J. 1996. Sus-a quick and dirty usability scale. usability evaluation in industry. 189(194):4–7.

Chai, J.; Ge, Y.; Liu, Y.; Li, W.; Zhou, L.; Yao, L.; and Sun, X. 2014. Application of frontal eeg asymmetry to user experience research. volume 8532 of *LNCS*, 234–243. EPCE 2014. Springer, Heidelberg.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn. Hillsdale, New Jersey: L.

Comrey, A. L., and Lee, H. B. 1992. *A first course in factor analysis*. Hillsdale, NJ: Erlbaum.

Devellis, R. F. 2003. *Scale Development: Theory and Applications,Second Edition, Applied Social Research Methods Series*, volume 26. Thousand Oaks, Calif.: Sage.

Fernandez, L.; Aulagnier, M.; Bonnet, A.; Guinard, A.; Pedinielli, J. L.; and Préau, M. 2005. *Module VI : Outils Psychométriques. Démarches épidémiologiques après une catastrophe*. Verger, P.; Aulagnier, M.; Schwœbel, V.; Lang, T.(Dir).

Georges, V.; Courtemanche, F.; Senecal, S.; Baccino, T.; Fredette, M.; and Leger, P. M. 2016. Ux heatmaps: Mapping user experience on visual interfaces. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 4850–4860. CHI2016. ACM.

Hassenzahl, M.; Burmester, M.; and Koller, F. 2003. *AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualitt [AttracDiff: A questionnaire to measure perceived hedonic and pragmatic quality]*. Ziegler, J.; Szwillus, G. (Eds.), Mensch & Computer 2003. Stuttgart, Leipzig: B. G. Teubner.

Heutte, J., and Fenouillet, F. 2015. Propositions pour une mesure de lexprience optimale (tat de flow) en contexte ducatif. In *Actes du 26e congrs international dactualit de la recherche en ducation et en formation (AREF)*. AREF2010.

Kennedy, R. S.; Lane, N. E.; Berbaum, K. S.; and Lilienthal, M. G. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology* 3(3):203–220.

Kim, C. J., and Chang, M. H. 2015. Actual emotion and false emotion classification by physiological signal. In *2015 8th International Conference on Signal Processing, Image Processing and Pattern Recognition (SIP)*, 21–24. SIP2015. IEEE.

Lewis, J. R., and Sauro, J. 2009. The factor structure of the system usability scale. *Human centered design* 94–103.

MacCallum, R. C.; Widaman, K. F.; Zhang, S.; and Hong, S. 1999. Sample size in factor analysis. *Psychological methods* 4(1):84–99.

McFarland, D. J.; Sarnacki, W. A.; Vaughan, T. M.; and Wolpaw, J. R. 2005. Brain-computer interface (bci) operation: signal and noise during early training sessions. *Clinical Neurophysiology* 116:56–62.

Murphy, C. A.; Coover, C.; and Owen, S. V. 1989. Development and validation of the computer self-efficacy scale. *Educational and Psychological measurement* 49(4):893–899.

Nunnally, J. C. 1978. *Psychometric Theory, (2nd ed.)*. McGraw-Hill Companies, 1978.

Oculus©. 2016. Oculus best practices. `https://developer.oculus.com/documentation/intro-vr/latest/concepts/bp_intro/`.

Pallot, M.; Eynard, R.; Poussard, B.; Christmann, O.; and Richir, S. 2013. Augmented sport: exploring collective user experience. In *Proceedings of the Virtual Reality International Conference: Laval Virtual (VRIC '13)*, 8. VRIC2013. ACM.

Pekrun, R.; Goetz, T.; Frenzel, A. C.; Barchfeld, P.; and Perry, R. P. 2011. Measuring emotions in students learning and performance: The achievement emotions questionnaire (aeq). *Contemporary Educational Psychology* 36(1):36–48.

Tabachnik, B. G., and Fidell, S. L. 2007. Discriminant analysis. using multivariate statistics. *Boston: Pearson Education Inc* 201(3):377–438.

Tcha-Tokey, K.; Loup-Escande, E.; Christmann, O.; Canac, G.; Farin, F.; and Richir, S. 2015. Vers un modèle de l'expérience utilisateur en environnement virtuel immersif : Une analyse de la littérature [towards a user experience in immersive virtual environment model: a review]. In *Proceedings of the 27th Conference on Interaction Homme-Machine*, 26. Toulouse, France: IHM2015. ACM.

Tcha-Tokey, K.; Loup-Escande, E.; Christmann, O.; and Richir, S. 2016. A questionnaire to measure the user experience in immersive virtual environment. In *Proceedings of the 18th VRIC 2016*. VRIC2016. ACM.

Venkatesh, V.; Morris, M. G.; Davis, G. B.; and Davis, F. D. 2003. User acceptance of information technology: Toward a unified view. *MIS quarterly* 27(3):425–478.

Witmer, B. G., and Singer, M. J. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence* 7(3):225–240.

Yao, L.; Liu, Y.; Li, W.; Zhou, L.; Ge, Y.; Chai, J.; and Sun, X. 2014. Using physiological measures to evaluate user experience of mobile applications. *In Engineering Psychology and Cognitive Ergonomics* 301–310.

Katy Tcha-Tokey completed her Human-Computer-Interaction (HCI) Master's degree diploma in the University of Toulouse III in 2013. She joined the Presence&Innovation (P&I) team in the LAMPA Laboratory in 2015 where she is pursuing a phD in the field of user experience in immersive virtual environment. Her current research interests include user experience modelling and instrument design for user experience validation.

Olivier Christmann is an associate professor in the LAMPA EA 1427 research laboratory belonging to Arts et Métiers ParisTech of Laval, France. He is director of a Master's course in Virtual Reality and Innovation. His current researches are focused on VR/AR and emerging technologies. He is interested in natural user interaction and in user experience.

Emilie Loup-Escande is an associate professor in the CRP-CPO EA 7273 research center of Picardy Jules Verne University in France. She is interested in the ergonomic design of emerging technologies. She has been involved in the design and assessment of virtual reality applications (product design, training and learning, entertainment).

Simon Richir, M.Eng, Ph.D., one of the pioneers and the most recognized leaders in French VR research and its practical application. Professor at ENSAM (Arts & Metiers ParisTech), the renowned French School of Engineering, Simon Richir is also the head of "Presence & innovation" research team (LAMPA Lab, EA1427). His research and teaching activities concentrate on technological innovation, engineering design process, innovative projects, and innovative uses of new advanced technologies such as Virtual Reality or Augmented Reality. In addition to these activities, SR is also the co-founder and the present scientific chair of one of the world's most prestigious international events in Virtual Reality the annual Laval Virtual International Conference.

APPENDIX 1

French version of our unified UX in IVE questionnaire (IVEQ) v2.

No v2	Items	Subscales	No v1
1	Mes interactions avec l'environnement virtuel me semblaient naturelles.	Présence	2
2	Les aspects visuels de l'environnement virtuel m'invitaient à m'y impliquer.	Engagement	3
3	Les périphériques (manette et/ou clavier) permettant mon mouvement dans l'environnement virtuel me semblaient naturels.	Présence	4
4	J'étais en mesure d'explorer activement l'environnement virtuel de faon visuelle.	Présence	5
5	La sensation de déplacement à l'intérieur de l'environnement virtuel était réaliste.	Engagement	6
6	À une distance proche, je pouvais bien examiner les objets.	Présence	7
7	Je pouvais examiner les objets sous différents angles.	Présence	8
8	J'étais impliqué(e) dans l'expérience vécue dans l'environnement virtuel.	Engagement	9
9	En termes d'interactions et de déplacements dans l'environnement virtuel, je me suis senti(e) compétent(e) à la fin de l'expérience.	Présence	10
10	Je suis parvenu(e) à me concentrer sur les tches requises plutôt que sur les périphériques d'interaction (manette et/ou clavier).	Présence	13
11	J'arrivais à identifier correctement les sons produits dans l'environnement virtuel.	Présence	14
12	J'arrivais à localiser correctement les sons produits dans l'environnement virtuel.	Présence	15
13	Je me sentais stimulé(e) par l'environnement virtuel.	Immersion	16
14	Il m'est arrivé d'être tellement absorbé(e) dans l'environnement virtuel que je n'étais plus conscient(e) des choses qui pouvaient se passer autour de moi.	Immersion	17
15	Il m'est arrivé d'être tellement absorbé(e) dans l'environnement virtuel que j'avais l'impression d'être à l'intérieur du jeu plutôt qu'en train de manuvrer des manettes de jeu et de regarder un écran.	Immersion	19
16	Physiquement, je me sentais bien dans l'environnement virtuel.	Immersion	20
17	Il m'est arrivé(e) d'être absorbé(e) dans le jeu à un point tel que j'en ai perdu la notion du temps.	Immersion	22
18	Je sentais que je contrôlais parfaitement mes actions.	Flow	23
19	À chaque étape, je savais ce que je devais faire.	Flow	24
20	J'avais le sentiment de bien contrôler la situation.	Flow	25
21	Le temps semblait s'écouler de faon différente de d'habitude.	Flow	26
22	J'avais l'impression que le temps passait rapidement.	Flow	27
23	Je perdais la notion du temps.	Flow	28
24	Je n'étais pas préoccupé(e) par ce que les autres pouvaient penser de moi.	Flow	30
25	J'avais le sentiment de vivre un moment enthousiasmant.	Flow	31
26	Cette expérience dans l'environnement virtuel me procurait beaucoup de bienêtre.	Flow	32
27	Quand j'évoque cette expérience dans l'environnement virtuel, je ressens une émotion que j'ai envie de partager.	Flow	33
28	J'ai apprécié être dans cet environnement virtuel.	Emotion	37
29	C'était tellement excitant à un point tel que je pourrais rester dans l'environnement virtuel pendant des heures.	Emotion	39
30	J'ai apprécié l'expérience à un point tel que je me sens plein(e) d'énergie.	Emotion	40
31	Je me suis senti(e) nerveux(se) dans l'environnement virtuel.	Emotion	41
32	J'avais envie de me distraire afin de réduire mon anxiété.	Emotion	44
33	Pendant que j'étais dans l'environnement virtuel, j'ai perdu le fil de mes pensées.	Emotion	45
34	Les périphériques d'interaction (casque Oculus, manette et/ou clavier) étaient très ennuyeux.	Emotion	46
35	Lorsque mes actions fonctionnaient, je ressentais une montée d'émotion positive.	Emotion	47
36	Lorsque j'utilisais ces périphériques d'interaction (casque Oculus, manette et/ou clavier), je trouvais le temps long.	Emotion	48
37	J'ai apprécié le défi que représentait l'apprentissage des périphériques d'interaction (casque Oculus, manette et/ou clavier) de réalité virtuelle.	Emotion	49
38	J'ai apprécié avoir affaire aux périphériques d'interaction (casque Oculus, manette et/ou clavier) de l'environnement virtuel.	Emotion	51
39	J'étais sûr(e) de moi lorsque je sélectionnais des objets dans l'environnement virtuel.	Compétence	52
40	J'étais sur(e) de moi lorsque je déplaais le viseur dans l'environnement virtuel.	Compétence	53
41	J'étais sûr(e) de moi lorsque j'utilisais la manette et/ou le clavier pour me déplacer dans l'environnement.	Compétence	54

Continued: French version of our unified UX in IVE questionnaire (IVEQ) v2.

No v2	Items	Subscales	No v1
42	Je suis sûr(e) de pouvoir comprendre les termes/mots relatifs aux périphériques d'interaction (casque Oculus, manette et/ou clavier) de l'environnement virtuel.	Compétence	55
43	Je suis sûr(e) de pouvoir acquérir des compétences avancées dans une application de réalité virtuelle utilisant le casque Oculus.	Compétence	56
44	Je suis sûr(e) de pouvoir décrire le fonctionnement des périphériques d'interaction (casque Oculus, manette et/ou clavier, joystick, kinect,) d'un environnement de réalité virtuelle.	Compétence	57
45	Personnellement, je dirais que cet environnement virtuel est Pas pratique/ Pratique.	Jugement	58A
46	Personnellement, je dirais que cet environnement virtuel est Confus/ Clair.	Jugement	58B
47	Personnellement, je dirais que cet environnement virtuel est Incontrôlable/ Matrisable.	Jugement	58C
48	J'ai trouvé que cet environnement virtuel était Ennuyeux/ Enthousiasment.	Jugement	59B
49	J'ai trouvé cet environnement virtuel Amateur/ Professionnel	Jugement	60A
50	J'ai trouvé cet environnement virtuel De mauvais goût/ De bon goût	Jugement	60B
51	J'ai trouvé cet environnement virtuel Non présentable/ Présentable	Jugement	60C
52	J'ai trouvé que cet environnement virtuel est Laid/ Beau	Jugement	61A
53	J'ai trouvé que cet environnement virtuel est Désagréable/ Agréable	Jugement	61B
54	Je me suis senti(e) fatigué(e) durant mon interaction avec l'environnement virtuel.	Conséquence de l'expérience	62
55	J'ai ressenti des maux de tête durant mon interaction avec l'environnement virtuel.	Conséquence de l'expérience	63
56	J'ai ressenti une fatigue visuelle durant mon interaction avec l'environnement virtuel.	Conséquence de l'expérience	64
57	J'ai remarqué une augmentation de ma salivation durant mon interaction avec l'environnement virtuel.	Conséquence de l'expérience	65
58	J'ai eu des nausées dans l'environnement virtuel.	Conséquence de l'expérience	67
59	J'ai eu une impression de lourdeur dans la tête durant mon interaction avec l'environnement virtuel.	Conséquence de l'expérience	68
60	Je me suis senti(e) étourdi(e) lorsque j'ouvrais les yeux durant mon interaction avec l'environnement virtuel.	Conséquence de l'expérience	69
61	J'ai ressenti des vertiges durant mon interaction avec l'environnement virtuel.	Conséquence de l'expérience	70
62	Si j'utilise de nouveau le même environnement virtuel, mon interaction avec l'environnement sera claire et compréhensible pour moi.	Adoption de la technologie	71
63	J'arriverai aisément à acquérir les compétences requises pour utiliser cet environnement virtuel.	Adoption de la technologie	72
64	Faire fonctionner à nouveau cet environnement virtuel sera facile pour moi.	Adoption de la technologie	73
65	Utiliser ces périphériques d'interaction (casque Oculus, manette et/ou clavier) de réalité virtuelle est une mauvaise idée.	Adoption de la technologie	74
66	Les périphériques d'interaction (casque Oculus, manette et/ou clavier) de réalité virtuelle pourrait rendre mon travail plus intéressant.	Adoption de la technologie	75
67	J'aimerais bien utiliser ces périphériques d'interaction (casque Oculus, manette et/ou clavier) de réalité virtuelle dans mon travail.	Adoption de la technologie	76
68	J'ai les ressources nécessaires pour utiliser ces périphériques d'interaction (casque Oculus, manette et/ou clavier) de réalité virtuelle.	Adoption de la technologie	77
69	À votre avis, quels ont été les points positifs de votre expérience ?	None	80
70	À votre avis, quels ont été les points négatifs de votre expérience ?	None	81
71	Avez-vous des suggestions pour améliorer cet environnement de réalité virtuelle ?	None	82

English translation of our unified UX in IVE questionnaire (IVEQ) v2 (originally in French).

No v2	Items	Subscales	No v1
1	My interactions with the virtual environment seemed natural.	Presence	2
2	The visual aspects of the virtual environment involved me.	Engagement	3
3	The devices (gamepad or keyboard) which controlled my movement in the virtual environment seemed natural.	Presence	4
4	I was able to actively survey the virtual environment using vision.	Presence	5
5	The sense of moving around inside the virtual environment was compelling.	Engagement	6
6	I was able to examine objects closely.	Presence	7
7	I could examine objects from multiple viewpoints.	Presence	8
8	I was involved in the virtual environment experience.	Engagement	9
9	I felt proficient in moving and interacting with the virtual environment at the end of the experience.	Presence	10
10	I could concentrate on the assigned tasks rather than on the devices (gamepad or keyboard).	Presence	13
11	I correctly identified sounds produced by the virtual environment.	Presence	14
12	I correctly localized sounds produced by the virtual environment.	Presence	15
13	I felt stimulated by the virtual environment.	Immersion	16
14	I become so involved in the virtual environment that I was not aware of things happening around me.	Immersion	17
15	I become so involved in the virtual environment that it is if I was inside the game rather than manipulating a gamepad and watching a screen.	Immersion	19
16	I felt physically fit in the virtual environment.	Immersion	20
17	I become so involved in the virtual environment that I lose all track of time.	Immersion	22
18	I felt I could perfectly control my actions.	Flow	23
19	At each step, I knew what to do.	Flow	24
20	I felt I controlled the situation.	Flow	25
21	Time seemed to flow differently than usual.	Flow	26
22	Time seemed to speed up.	Flow	27
23	I was losing the sense of time.	Flow	28
24	I was not worried about what other people would think of me.	Flow	30
25	I felt I was experiencing an exciting moment.	Flow	31
26	This experience was giving me a great sense of well-being	Flow	32
27	When I mention the experience in the virtual environment, I feel emotions I would like to share.	Flow	33
28	I enjoyed being in this virtual environment.	Emotion	37
29	It was so exciting that I could stay in the virtual environment for hours.	Emotion	39
30	I enjoyed the experience so much that I feel energized.	Emotion	40
31	I felt nervous in the virtual environment.	Emotion	41
32	I felt like distracting myself in order to reduce my anxiety.	Emotion	44
33	I found my mind wandering while I was in the virtual environment.	Emotion	45
34	The interaction devices (Oculus headset, gamepad and/or keyboard) bored me to death.	Emotion	46
35	When my actions were going well, it gave me a rush.	Emotion	47
36	While using the interaction devices (Oculus headset, gamepad and/or keyboard), I felt like time was dragging.	Emotion	48
37	I enjoyed the challenge of learning the virtual reality interaction devices (Oculus headset, gamepad and/or keyboard).	Emotion	49
38	I enjoyed dealing with the interaction devices (Oculus headset, gamepad and/or keyboard).	Emotion	51
39	I felt confident selecting objects in the virtual environment.	Skill	52
40	I felt confident moving the cross hair around the virtual environment.	Skill	53
41	I felt confident using the gamepad and/or keyboard to move around the virtual environment.	Skill	54
42	I feel confident understanding the terms/words relating to the interaction devices (Oculus headset, gamepad and/or keyboard).	Skill	55
43	I feel confident learning advanced skills within a specific virtual reality software using the Oculus headset.	Skill	56
44	I feel confident describing the functions the interaction devices (Oculus headset, gamepad and/or keyboard, joystick, Kinect . . .) of a virtual reality environment.	Skill	57
45	Personally, I would say the virtual environment is impractical/practical.	Judgement	58A

Continued: English translation of our unified UX in IVE questionnaire (IVEQ) v2 (originally in French).

No v2	Items	Subscales	No v1
46	Personally, I would say the virtual environment is confusing/clear.	Judgement	58B
47	Personally, I would say the virtual environment is unruly/manageable.	Judgement	58C
48	I found that this virtual environment was lame/exciting.	Judgement	59B
49	I found this virtual environment amateurish/professional.	Judgement	60A
50	I found this virtual environment gaudy/classy.	Judgement	60B
51	I found this virtual environment unpresentable/presentable.	Judgement	60C
52	I found that this virtual environment is ugly/beautiful.	Judgement	61A
53	I found that this virtual environment is disagreeable/likeable.	Judgement	61B
54	I suffered from fatigue during my interaction with the virtual environment.	Experience consequence	62
55	I suffered from headache during my interaction with the virtual environment.	Experience consequence	63
56	I suffered from eyestrain during my interaction with the virtual environment.	Experience consequence	64
57	I felt an increase of my salivation during my interaction with the virtual environment.	Experience consequence	65
58	I suffered from nausea during my interaction with the virtual environment.	Experience consequence	67
59	I suffered from fullness of the head during my interaction with the virtual environment.	Experience consequence	68
60	I suffered from dizziness with eye open during my interaction with the virtual environment.	Experience consequence	69
61	I suffered from vertigo during my interaction with the virtual environment.	Experience consequence	70
62	If I use again the same virtual environment, my interaction with the environment would be clear and understandable for me.	Technology adoption	71
63	It would be easy for me to become skillful at using the virtual environment.	Technology adoption	72
64	Learning to operate the virtual environment would be easy for me.	Technology adoption	73
65	Using the interaction devices (Oculus headset, gamepad and/or keyboard) is a bad idea.	Technology adoption	74
66	The interaction devices (Oculus headset, gamepad and/or keyboard) would make work more interesting.	Technology adoption	75
67	I would like working with the interaction devices (Oculus headset, gamepad and/or keyboard).	Technology adoption	76
68	I have the resources necessary to use the interaction devices (Oculus headset, gamepad and/or keyboard).	Technology adoption	77
69	In your opinion, what were the positive points about your experience?	None	80
70	In your opinion, what were the negative points about your experience?	None	81
71	Do you have suggestions to improve this virtual reality environment?	None	82