Impact of Verbal Communication on User Experience in 3D Immersive Virtual Environments

Rémy EYNARD, Marc PALLOT, Olivier CHRISTMANN, Simon RICHIR
Arts et Métiers ParisTech, LAMPA
2, boulevard du Ronceray
49000 Angers, France
remy.eynard@ensam.eu, marc.pallot@ensam.eu, olivier.christmann@ensam.eu, simon.richir@ensam.eu

***Abstract*—Understanding the user experience has become over the last decade a major concern for both scientific community and industry. In the field of virtual reality, user experience models have already been established, however, only some of them have been empirically validated. Furthermore, the potential impact of verbal communication on user experience appears to be marginally studied by scholars. This study aims to assess the impact of verbal communication on specific factors of the user experience in the context of Virtual Reality. These factors are spatial presence, social presence and performance gain. Participants (N = 52) immersed by teams of two players in the game Minecraft had to achieve a specified goal in a given period of time. To measure the impact of verbal communication on this task by comparison, half of teams could not communicate verbally, the other half were able to. Our results highlight a significant positive effect of verbal communication on the performance gain as well as on the hedonic evaluation of the user experience. Despite some limitations and questions raised by this empirical study, these results along with previous literature demonstrate the significant impact induced by interpersonal verbal communication on the user experience in the context of a 3D immersive virtual environment.**

Keywords— Virtual Reality ; 3D Immersive Environment ; User Experience ; Social Interactions ; Verbal Communication ; Spatial Presence ; Social Presence ; Performance

I. INTRODUCTION

Virtual reality (VR) has become in recent years increasingly common and accepted, thanks to the vulgarization of display system, such as the Oculus Rift[1]. This fast evolution opens new areas of exploration for VR uses, for example in the field of psychology, which demonstrated a growing interest in VR technologies. Indeed virtual environments (VEs) offer to psychologists a situational control absolutely unattainable in reality. Moreover some studies assess that social attitudes shown by users in VEs are highly correlated with their behaviors and personality in the real world [1]. One of the factors that contributed to the growing interest of the scientific community with respect to mediated social behaviors (through a media, screen, phone, video game) is the explosion of games or virtual platforms providing a persistent virtual universe and giving an important place to collaboration and interaction between players. In addition to verbal communication (VC) we could see the apparition of multi-player gameplay based almost solely on Non-Verbal Communication (NVC) [2]. Researchers show a strong interest in these platforms which are privileged areas for studies about users behaviors in virtual and connected environments because of their popularity and of the engagement they arouse [3]. These applications require a component having a high complexity in its virtual reproduction that arouses great interest during the last decade: social interaction [4].

Socially interacting with an agent[2] or an avatar[3] in a natural way is not obvious. A lot of components are involved to create a coherent social interaction: the voice, its intonation and volume, the face of the speaker, his expressions, his color (if it turns red for example), body, hands movements and posture [5]. These elements are inseparable of a viable social interaction and ask for a trustful representation in VR in order to not appear uncanny to users [6].

The user experience (UX), i.e. the whole set of perceptions and responses of a person regarding to the use or the anticipation of the use of a product or service, is now widely used to evaluate and so to design or make evolve such products and services. Many definitions of UX exist in the literature [7]. Crossing these definitions provides a global view of factors compounding UX and especially when this one is applied to 3D Immersive Virtual Environments (3D-IVEs). Wu [8] uses the terms Distributed Interactive Multimedia Environments (DIMEs), but we prefer in our context the use of 3D-IVE as described for example by Garau et al. [9]. In this study, we use primarily the work of Wu et al. [8], because of its really relevant dichotomy of UX in two parts: Quality of Service (QoS) and Quality of Experience (QoE). Then, we rely on the work of Pallot et al. [10] who propose an enriched version of Wu's model.

The main objectives of the study are: (a) Obtain measures of UX factors in order to (b) have a better understanding of the impact of VC on UX factors during the immersion in a 3D-IVE. By a better understanding of eventual correlations between social interaction and UX factors [11], we would like to facilitate the creation of applications oriented on social interaction (for example dedicated to distant sports, works, conferences, relationships). Through this paper we try to validate the following hypothesis: the ability to interact verbally with a partner, when immersed in a 3D-IVE, impacts factors of UX and thus improves the perceived QoE. Increasing QoE and understanding impacting factors would allow to significantly improving current distant working technologies. The tight integration of emotional and emphatic aspects, knowing that they have a strong impact on presence [12], allows to improve: performance, mutual understanding and awareness in the context of telepresence.

1 The Oculus Rift is an HMD (Head Mounted Device), please see: http://www.oculus.com

2 Virtual Agent: A virtual representation of an intelligent being controlled by a computer.

3 Virtual Avatar: A virtual representation of an intelligent being controlled by a human.

II. EXISTING THEORIES & PREVIOUS WORKS

A. Social interaction

Regarding social interaction, we can distinguish two kinds of communication: Verbal Communication (VC) and Non-Verbal Communication (NVC) (e.g. body language) [5]. In the VC, we can distinguish several subsets. First, the prosodic intonation helps to provide an expressive modality to the message. Second, there are two fairly simple factors: speech speed (or flow) and sound volume. They complement intonation by giving information about the intensity of the feeling. These three elements (intonation, volume and flow), combined in various shapes, allow a wide variety of modulation on the same sentence. In the scope of VR, they are frequently used to assess the level of effectiveness of a person regarding to a task involving social interaction. We can see an illustration through the creation of projects such as "MACH: My Automated Conversation Coach" that helps participants to improve their effectiveness in the context of job interviews [13].

NVC can also be divided into several elements, in the same way as for the VC. These factors are part of what we call social behavior [5]. Posture is the broader element because it refers to a general view of the person (transmitter or receiver of the message): it can of course provide information about health status or age of a person, but also as shown by Pertaub et al. through an experiment based on the presentation of a speech to an audience [14], the disposition to listen and the interest paid to the received message. Pertaub et al. show in their study that a plausible and credible social interaction is possible in a 3D-IVE [14]. Implementation and realism of gestures [15], facial expressions [6], mutual gaze [16] and proxemics [17], are essential to generate sense of presence in 3D-IVEs.

B. User Experience

Since the introduction of user-centered design methodologies, UX has become very popular [7]. As defined by ISO 9241-210: "UX is perceptions and responses of a person resulting from the use or the anticipation of the use of a product or a service". Despite the fact that many different models exist, Law et al. [18] have shown that there is a consensus on the dynamic, subjective and context-dependent nature of UX. This is partly explained by the fact that a person is carrying his unique past experiences, which may alter, influence new experience but also increase the level of expectation with respect to the next occurrence [19].

In the scope of our study we particularly focus on UX models related to both QoS and QoE. The later one, as described by Gaggioli et al. [20], is the subjective part of UX, its main component is the sense of presence also called telepresence or spatial presence, synthetized as "The sense of being there" in a 3D-IVE. On the other hand QoS corresponds to the measurement of technical performances [21]. The UX model proposed by Wu et al. correlates QoS and QoE [8] through a causal relationship. In addition, Wu proposes a framework highlighting factors involved in the causal chain. Pallot [10] adds to Wu's framework the concept of social interactions and interpersonal communication in a DIME.

C. Measuring User Experience

Inspired by the theory of GameFlow [22], we have seen emerge the Presence-Flow framework, divided in three dimensions: Physical presence, Situational involvement and Competence. Commonalities among these theories of definition of UX are:

- The division into factors. From technical factors (QoS) to the most subjective human factors (QoE).
- The measurement of these factors in order to assess UX.
- The consensus saying that a good UX is the result of the balance between its factors. There is no single factor on which one to act but an interdependent whole.

1) User experience measurable factors:

Based upon models defined by Wu and Pallot, we aim to evaluate several factors of UX, namely: telepresence, performance gain, emphatic and emotional behaviors. This choice was made regarding the ability to measure these factors but mostly with respect to the fact that this study compares two states of social interactions. The state of Social Presence (ScP) allows the user the use of VC during his immersion in the 3D-IVE, contrary to the other state called Co-Presence (CP). Assessed factors had to be measurable in the same way in these two states to allow their comparison. The following paragraphs explain briefly the nature of factors (for a complete review, see [10]) but especially the way we measure them.

a) Immersive (U)X > Rational Engine > Telepresence:

Telepresence, spatial presence or simply presence represent for users the feeling of "being there" in the VE. We use the term Spatial Presence (SP) to avoid confusion with Social Presence (ScP). At the basis of this concept Lombard and Ditton [23] postulate the need of a certain realism to achieve the sense of presence, which has been confirmed by Bailenson et al. [6].

At the same period Heeter [24] proposed a vision more suited to 3D-IVEs and divided into three parts. First, personal presence: the fact that the person feels immersed in the environment, as an integral part of it, that is SP. Second, ScP: the feeling of presence of other beings in the environment whether incarnated by a human or an avatar or virtual agent. Third, environmental presence: the fact that the environment itself is "alive" and responds to actions / interactions of the immersed person. This part is linked to the "Interactivity" of the QoS block in Wu's model [8]. Despite the many visions sometimes contradictory of the concept of SP, a consensus seems now established: the spatial presence is a subjective phenomenon, a feeling ("being there"). The most common SP measure remains a subjective assessment using a Likert items based questionnaire. Those developed by Schubert [25] or Freeman [26], are reliable in the evaluation of SP in a 3D-IVEs.

b) Immersive (U)X > Experiential Engine > Psychological Flow > Enjoyment:

Pallot describes the psychological flow as follows: "Psychological Flow represents the feelings of someone acting with total involvement procuring the perception of great enjoyment and sense of control. Activities such as reading, gaming or sporting provide an intense feeling of immersion as a natural flow of mind" [10]. Enjoyment is a hedonic feeling. Usage time, and the desire to reuse are good cues of the enjoyment felt during an experience. The easiest way to measure enjoyment remains questionnaire.

c) Immersive (U)X > Experiential Engine > Psychological Flow > Social presence:

Like SP ("being there"), ScP may be described as the feeling of being present with another entity ("being there with another"). In the taxonomy of presence proposed by Heeter

[24], the importance of ScP is stressed as follows: "The premise of ScP is simply that if other people are in the virtual world, that is more evidence that the world exists". Biocca & Harms have proposed a multimodal construction of ScP divided in three factors:

- **"Co-Presence:** The degree to which the observer believes he is not alone, his level of peripherally or focally awareness of the other, and their sense of the degree to which the other is peripherally or focally aware of them". We can say that Co-Presence is the *SP of the other*, the feeling of being present with another and to be able to interact with him in a direct or an indirect way (via the 3D-IVE). Social interactions that can then occur (if the VC is not possible) are about NVC. This observation of the presence of the other leads to an empathic behavior: translucence, which tends to establish an understanding relationship between two individuals [27].
- **"Psychological Involvement:** The degree to which the observer allocates focal attention to the other, empathically senses or responds to the emotional states of the other, and believes that he/she has in-sight into the intentions, motivation, and thoughts of the other."
- **"Behavioral engagement:** The degree to which the observer believes his/her actions are interdependent, connected to, or responsive to one-another and the perceived responsiveness of the other to the observer's actions" [28]. Durlach et al. had already demonstrated empirically that a shared interaction between participants in 3D-IVEs had a positive impact on the feeling of ScP.

There are also other factors impacting the ScP particularly levels of anthropomorphism and realism of a virtual character (visual and behavioral) [6], [15]. Nevertheless, we can say in accordance with UX models from Wu & Pallot, that this factor is related to the QoS rather than to the QoE. It has an influence on ScP but if it does not vary, it can allow a comparison between two states, which is our objective. We distinguish subjective and objective measures.

The ScP is objectively measurable mainly through observation. Speaking time in the context of social interactions through the VC is a good indicator of social engagement [29] (corresponding to the behavioral engagement described by Biocca). We can also observe the intensity (volume of the voice and its variations) of these social interactions by the VC. We can also attempt to identify the nature of VCs, through the tone, the speech rate, or used terms. It is also possible, as argued by Biocca et al. [30], to observe the following behaviors, such as: attentional behaviors such as eye fixation on the one-another and proxemics behavior, movement to or away.

d) QoE > Empathical Behaviour:

Empathical behaviours are behavioral responses of a subject in CP or ScP, it can be encouragements, help proposals, spontaneous assistance or advices. These emphatical behaviors are generators of social emotions such as embarrassment, guilt, shame and pride. These emotions are difficult to measure objectively, so their assessment rely on questionnaire or observation.

e) QoE > Emotionnal Behaviour:

Emotional behaviour can arise both in a single or multi-user activity, they represent the personal emotional response of a subject to an activity, a stimulus, whether generated by the 3D-IVE or by a partner in ScP. These emotional behaviours are measurable through questionnaire or observation.

f) QoE > Performance Gain:

The performance gain represents the increase of individual performance of a user according to both hedonic (pleasure, joy) and ergonomic (effort) perspectives. This factor can be objectively measured by an observation of the action progress, through the completion time, its success, how it is done or the number of mistakes. As in UX models proposed by Wu and Pallot, the hypothesis of a positive correlation between cognitive experience and performance gain is emitted.

Our goal is to reach a better understanding of the impact of social interaction on UX. Through this paper, we try to validate the following hypothesis: *The ability to interact verbally with a partner, when immersed in a 3D-IVE, impacts positively UX in a way that it increases the perceived QoE.*

III. RESEARCH METHODS

A. Subjects

Participants (N = 52; 13 female, 39 male) were aged from 19 to 27 years old (M = 21.15). We chose an homogenous population because the age of subjects has an impact on the level of ScP experienced [31]. All participants were fluent French speakers. They all signed up voluntarily to the experiment. Only two persons said that they were not video game players at all, the rest of the participants play video games at different frequencies (59.6% of the participants play less than 5 hours per week, 25% play from 5 to 15 hours per week and only 6% play more than 15 hours per week). In order to be sure that most of the participants were in ease with the use of the environment and gameplay modalities we also asked them on which platforms they play: most of them play on computers (84,6%), and less on home consoles (36,5%) but all the participants knew how to use a gamepad.

B. Subjects equipment and related

In order to achieve a high level of immersion, we immersed participant through the Oculus Rift DK1. We used the game Minecraft, and to make it usable on the Oculus Rift we used a mod[4] called Minecrift[5].

Participants were equipped with headphones. Finally, we used a webcam microphone and the software TeamSpeak[6].

As participants wearied an Oculus Rift their view of the real world was totally obstructed. Hence, we decided to use a gamepad (Microsoft Xbox 360 wired gamepad), which is easily usable even with eyes closed (indeed when one plays with a gamepad, he almost never looks at it).

C. Virtual environment

The current study use Minecraft (Markus Pearson, Mojang Inc., version 1.6.2 for MS Windows). Minecraft is a type of game called "sand box" that means the game offers a high degree of liberty to users regarding to their actions. Players have a large range of craftable objects that need different materials that needs to be mine. We focused on the creative game mode, in which the player can get access to every resources of the game in an unlimited way to create buildings or objects. However, its popularity, its ease of use, multiplayer gameplay and game mechanics has motivated our choice to select this game for this experiment.

4 Mod: Additional content, script or program developed by the player community for a video game.

5 Minecrift's website: http://minecraft-vr.com

6 TeamSpeak's website: http://www.teamspeak.com

In the context of this experiment the default textures pack of Minecraft was judged a little bit aggressive for participant's immersed trough the Occulus Rift. Because realism is an important factor of immersion [32], we decided to try to increase the realism level by changing some textures using the textures pack called PureBDCraft[7]. In order to save computer resources we used its 64 pixels texture version.

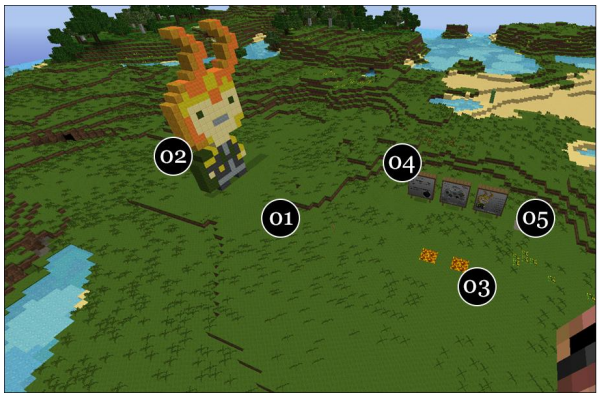

Fig. 1. Experiment area in the Minecraft virtual environment.

For this experiment a world was generated and modified in order to fit to the requirements. This Minecraft virtual world can be seen from two sides:

- **Experimenter side:** An area compound of special blocks allowing the controls of the area and of participants.
- **Participant side:** A simple area filled with all elements needed to perform a challenge. The area is composed of five elements (see Figure 1.): Bubble 01 is the place where participants start, this point is placed between the real size model of the statue (bubble 02) that they will have to reproduce; bubble 03 is the position landmarks placed to help them building their first tiers correctly; bubble 04 represents three help boards: the first one summarizes the tasks to be completed, the second one is a map of the gamepad control, and the last one is a construction plan of the statue; bubble 05 corresponds to an item dispenser dedicated to the second part of the experiment that is explained in the next paragraph.

D. Procedure

We have ensured that participants cannot communicate with each other before the experiment. No information was given about the real objective of the experiment to the participants in order to avoid any potential influence on either their answers to the questionnaires or their state of mind. Before equipping participants, we asked them to fill a quick pre-experiment questionnaire to record information about their knowledge of Minecraft and their eventual predisposition to motion sickness. After that, participants were instructed about the evaluation questionnaires used for this experiment. Explanations were also provided about the Minecraft user interface: use of the inventory, items in the inventory and their use, aspect of avatars and some other game modalities.

The experiment was designed for pair of participants working together on a specific challenge. To be able to quantify the impact of VC on UX, we designed this experiment as a comparison between two states: one without VC that we call the *CP state* and the other one with VC that we call the *ScP state*. We separated randomly the participants in two groups to obtain 13 pairs for each state.

Participants had to perform two different tasks; the first one was designed to encourage participants to collaborate (duration 25 min). The second one aimed to obtain strong emotional feedbacks from participants (5 min). During the first task, participants had to recreate a statue, which was composed of different Minecraft materials. In order to help them, a real size model of the statue and a board with a detailed blueprint including a black edging for facilitating the reading of the map (ease for counting blocks), of the relationship between each color and the matching material in the participants' inventories, were present in the environment (see Figure 1). The second task was less complex, participants had to defend their statue from groups of monsters that appeared by waves and tried to destroy a block placed above the statue. At the end participants were asked to fill a post-experimentation questionnaire.

E. Measurements

We present in this section the measures dedicated to a better understanding of the impact of VC on UX in 3D-IVEs. For an easier reading, this section is divided according to the different subjects of measurement (blocks of the UX) and not regarding the measurement tools we used. We still rely, for QoE factors of UX, on the work of Wu [8] and Pallot [10].

1) Spatial Presence

(Immersive (U)X > Rational Engine > Telepresence)

To measure the impact of the two types of session (CP & ScP) on SP, we established a five-level Likert item questionnaire in which the questions are extracted from the same "Igroup Presence Questionnaire" from [33] (see questions 1 to 5 of IPQ[8]).

2) Co-Presence

(Immersive (U)X > The rational Engine > Psychological Flow > ScP)

To measure the impact of the two types of session (CP & SCP) on the feeling of CP, we added 6 questions to our questionnaire. These questions are directly extracted from the questionnaire "Networked Minds Questionnaire" of Biocca & Harms [34] and evaluate the following factors to obtain a measure of the CP: Attentional allocation, Mutual understanding, Behavioral interdependence, Mutual assistance.

3) Performance Gain

(QoE > Performance Gain)

To measure the impact of the two types of session (CP & ScP) on the performance gain, we established two types of measures: one based on a subjective five-level Likert item questionnaire extracted from the previous work of Pallot, the second one extracted directly from the questionnaire "Networked Minds Questionnaire" of Biocca & Harms [34]. Additionally we introduced an evaluation based on the observation of participant actions to assess their level of performance with respect to the task to be completed. These observations consisted in counting the number of blocks placed at the end of the allocated time for the construction of the statue (25 minutes) and in checking if participants made or not construction mistakes.

8 Igroup Presence Questionnaire website: http://www.igroup.org/pq/ipq/index.php

4) Additional information on the assessment of social interactions of participants with the experience lived by ScP

(Immersive (U)X > The rational Engine > Psychological Flow > ScP & QoE > Empathical & Emotional Behavior)

We added to the questionnaire few questions dedicated to the ScP case in order to gather some clues about there VCs (frequency, intensity, nature of interactions).

5) A Posteriori questions: emotional memories and satisfaction

(Immersive (U)X > Experiential Engine > Psychological Flow > Enjoyment)

A five-level Likert item questionnaire with four questions (see below) was sent to the participants 4 months after the experiment to assess the emotional potential of this experience and their level of satisfaction from a hedonic point of view.

IV. FINDINGS

A. Data analysis

As a starting point, we conducted several tests of normality (Kolmogorov–Smirnov) on each variable obtained from the questionnaires in order to see if we could use parametric tests or not. None of those variables followed a normal distribution. These results are consistent with our data because Likert scales included few items. A better understanding and greater appropriation of the final questionnaire by participants motivate this choice. Consequently, we conducted tests belonging to the family of nonparametric tests. Given our experimental design (2 groups of participants x 2 conditions), we relied on the Mann-Whitney U-test for independent samples. It is from this test that we established the statistical significance of our results. We used the conventional threshold probability $p < 0.05$.

A set of answers to the questionnaire is considered valid if the two-person team was immersed at least 10 minutes. Indeed, participants had the time, during this period, to get used to the virtual environment, to interact with it and to communicate with each other (if the condition allowed it). The average time before laying the first block (45.42 seconds) and the average number of blocks set in 10 minutes (93.65 blocks on 200.52 blocks per pair over the duration of the experiment) confirms the validity of this choice.

B. What is the impact of VC on participant SP?

(Immersive (U)X > Rational Engine > Telepresence)

Five questions in our questionnaire measured the SP felt by participants during their experience. A Mann-Whitney U-test revealed a significant difference in the evaluation of the spatial presence regarding to the question: "You had a sense of acting in the virtual space, rather than operating something from outside" between the CP state (N = 26) and the ScP state (N = 26) of the experiment (U = 235, z = -2.009, p = .045). No significant difference was revealed regarding the other four questions in the spatial presence part of the questionnaire.

In order to know in which of the two states the feeling of spatial presence was best evaluated by participants, the average response was calculated for these two states. The following averages represent the average positioning of participants on the terms of the 5 points Likert scale: CP M = 4.38, ScP M = 3.85.

C. What is the impact of VC on CP?

(Immersive (U)X > The rational Engine > Psychological Flow > ScP)

Six questions in our questionnaire measured the CP experienced by participants through the previously described factors. A Mann-Whitney U Test revealed a significant difference in the evaluation of the CP regarding to the following factors:

1) Mutual understanding through the questions:

a) "The other person seemed to understand what you wanted to do (actions)": between the CP state (N = 26) and the ScP state (N = 26) of the experiment (U = 116, z = -4.172, p < .0001).

b) "I understood what the other person wanted to (actions)": between the CP state (N = 26) and the ScP state (N = 26) of the experiment (U = 156, z = -3.453, p = .001).

2) Behavioral interdependency through the question: "Your actions were dependent on those of the other person": between the CP state (N = 26) and the ScP state (N = 26) of the experiment (U = 160.5, z = -3.393, p = .001).

No significant difference was observed with respect to the other three questions of the CP part of the questionnaire. In order to know in which of the two states the CP was best evaluated by participants, the average response was calculated for both states: 1.a) CP M = 2.58, ScP M = 4.00 - 1.b) CP M = 2.88, ScP M = 3.96 - 2) CP M = 2.27, ScP M = 3.19.

D. What is the impact of VC on performance gain?

(QoE > Performance Gain)

1) Subjective measures of performance gain

Eleven questions in our questionnaire measured the performance gain felt by participants during the experience. A Mann-Whitney U Test revealed a significant difference in the evaluation of performance gain regarding to the followings question:

1) "You have been effective regarding to the completion of the task" between the CP state (N = 26) and the ScP state (N = 26) of the experiment (U = 207, z = -2.487, p = .013).

2) "You had the feeling of working with your partner on the task" between the CP state (N = 26) and the ScP state (N = 26) of the experiment (U = 152, z = -3.512, p < .0001).

3) "I had fun during this experience" between the CP state (N = 26) and the ScP state (N = 26) of the experiment (U = 235, z = -2.064, p = .039) (see Table 1).

	Question 1[1]	**Question 2[2]**	**Question 3[3]**
U of Mann-Whitney	207	152	235
Z	- 2.487	- 3.512	- 2.064
Asymp. sig.	.013	.000	.039

1 You have been effective regarding to the completion of the task.

2 You had the feeling of working with your partner on the task.

3 I had fun during this experience.

TABLE I. NON PARAMETRIC TEST ON PERFORMANCE GAINS PARTICIPANTS' EVALUATION.

No significant difference was observed with respect to the other eight issues of performance gain part of the questionnaire. In order to know in which of the two states the performance gain was best evaluated by participants, the average of the responses was calculated for each of the two states: 1) CP M = 2.42, ScP M = 3.15 - 2) PC M = 2.42, ScP M = 3.54 to 3) CP M = 4.00, ScP M = 4.54.

2) Objective measures of performance gain

These results are based on observation of the number of blocks set down by participants within 25 minutes and the fact that they made errors or not. To be sure of the robustness of this measure, we asked participants (in the identification questionnaire) what was their experience of the Minecraft game in order to test its influence on the performance gain. We also examined the distribution of participants in the groups of the experiment (CP & ScP) regarding their experience. Despite the randomness of the constitution of the pairs, we obtained almost a homogeneous distribution (see Table 2).

	No experience	**From 1 to 50h**	**More than 50h**	**Total**
Co-Presence	16	03	07	26
Social Presence	16	02	08	26

TABLE II. REPARTITION OF PARTICIPANTS' MINECRAFT EXPERIENCE IN THE TWO PARTICIPANTS GROUP.

In order to evaluate the possible interaction between the two session types (CP and ScP), blocks placed and the previous experience on Minecraft, we could not use a non-parametric test. ANOVA is required, although the distribution of our performance measurements deviates from the normality. We chose to perform an ANOVA, given its robustness for type 1 errors [35]. The two ways ANOVA did not reveal any significant interactions between the three factors mentioned above, we can now turn our attention to the main effects between factors i.e. the impact of Minecraft preliminary experience on the number of blocks placed and the impact of the session type (CP or ScP) on the number of blocks placed. A Kruskal-Wallis Test revealed no statistically significant difference in the the number of blocks placed regarding to pre-experiment Minecraft experience. A Mann-Whitney U-test revealed a significant difference in the number of blocks placed between the CP state and the ScP state of the experiment ($U = 188$, $z = -2.410$, $N=26$, $p = .016$).

In order to know in which of the two states the number of blocks placed was the highest, the average of blocks placed was calculated for the two state (CP M = 175.62, ScP M = 227.50). This reveals that participants who lived the experience on ScP state reach an average of 51.88 more blocks than those who lived the experience on the CP state. Regarding occuring errors, we observe that in a state of CP, 7 pairs have made mistakes. In a state of ScP, 2 pairs have made mistakes.

E. A Posteriori questions: emotional memories and satisfaction (Immersive (U)X > Experiential Engine > Psychological Flow > Enjoyment)

A Mann-Whitney U Test revealed a significant difference in the evaluation of the following question: "What was your feeling about this experience?" (From 1 = "Boring" to 5 = "Captivating") between the CP state (N = 18) and the ScP state (N = 12) of the experiment ($U = 60.500$, $z = -2.123$, $p = .043$).

	Emotional potential		**Satisfaction level**	
	Question 1[1]	**Question 2[2]**	**Question 3[3]**	**Question 4[4]**
Co-Presence	3.17 (1.150)[5]	3.61 (1.243)	3.22 (1.437)	3.17 (1.465)
Social Presence	3.75 (.965)	4.50 (.674)	3.50 (1.243)	3.92 (1.311)
Total	3.40 (1.102)	3.97 (1.129)	3.33 (1.348)	3.47 (1.432)

1 Which of these illustrations best describes the emotional state you felt during the experiment?
2 What was your feeling regarding to this experience?
3 Would you like to use again in a personal context this application?
4 Would you recommend the use of this application to others?
5 Mean (SD)

TABLE III. MEANS OF PARTICIPANTS' ANSWERS AT THE AT POSTERIORI QUESTIONNAIRE ABOUT EMOTIONAL POTENTIAL OF THE EXPERIENCE AND SATISFACTION LEVEL.

Despite the significant difference between the two states with respect to this question, this result should be considered as a trend because of the difference in number of subjects between the two states. Table 3 synthesizes these results.

V. DISCUSSION

Before the discussion of our results and their significance, we present the limitations of this work. First this experiment was designed to obtain joint subjective and objective measures with respect to the studied three factors, namely: SP, CP, performance gain, in order to confirm the consistency of the results; unfortunately, due to technical issues, we were unable to implement the planned physiological objective measures for the SP and CP factors. Hence, it has restricted our evaluation to questionnaires that are collecting subjective data.

In fact we can say that having VC in 3D-IVEs has an effect on the studied factors but not in what extent, and whether there are potential interactions between these factors. It would be desirable to obtain a graphic representation of the correlation including its intensity as done theoretically by Wu et al. [8]. To conclude, our results allow us to only partially confirm the impact of VC on UX, despite the reliability of questionnaire models and our observational measures. The results of this study therefore have to be confirmed.

Despite the discussed limitations, the results of this study confirm to some extent our main hypothesis. The ability to interact vocally during an immersion in a 3D-IVE impacts the user experience in a way that it improves the perceived QoE. These results are a step closer to a better understanding of levers that provide a high quality experience during the immersion in a 3D-IVE and consistent with the work of Wu [8] and Pallot [10]. The conclusions we can make from these results open up new fields of investigation. We observe a lower evaluation of the spatial presence by participants during the sessions where it was possible to communicate vocally. Does the act of speaking bring participants back to an awareness of the real world in a way that it has induced a "Break In Presence" (BIP) [36]?

Unfortunately, the lack of work on this BIP issue does not allow us to provide a definitive answer. We must also take into account the relatively low level of realism of the avatars in Minecraft. For example, avatars could not show facial expressions or even move the mouth during verbal interactions. For this reason, the voice can be perceived as not coming from the avatar, which does not stick to the representation of a normal voice interaction. This problem occurs most probably in the phenomenon of BIP [37]. We later found that during the ScP sessions, the presence of one-another (CP) and their mutual understanding were significantly higher. We also observe that participants immersed in a state of ScP have experienced greater behavioral interdependence: the opportunity to speak encourages participants to find the best strategy for efficiency. In this quest for efficiency, often one of the protagonists becomes the leader, which explains the above-mentioned phenomena that occurred during ScP sessions. It is certain that speech allowed participants to reduce the "difficulty" of the task, through collaboration, organization, strategy, leadership decision, sharing of abilities and skills of the participants. This observed collaboration behaviors allowed

them to be more efficient and make fewer mistakes. However, can we therefore deduct that having a voice exchange increases the performance gain in all cases?

For example, in the case of a complementary but not interdependent task, it seems logical that participants would reduce their VC. In contrast, in the case they persist to use VC by encouraging themselves, discussing about progresses and the state of mind of each other, we do anticipate that there would be an impact on the performance gain.

In addition to the above observations on the performance gain, we do assume that the nature of the task would have an impact as well on the perceived hedonic quality. Indeed we observed, through the analysis of the post-experimentation collected data, that participants engaged in ScP rated their experience as more exciting than their counterparts who were unable to communicate vocally. Therefore, participants have expressed, when engaged in ScP, a higher satisfaction and hedonism. However, if VC tends to rather disturb participants then one could assume that the impact could be different. Hence, this is something that we would like to address in a future experiment.

VI. Conclusion & Future work

According to our hypothesis, we expected to obtain significant results on the impact of verbal communication through the rating of spatial presence. This hypothesis was not validated, and we even got a slightly more positive assessment of this factor in the state of co-presence. However, the observed impacts on the performance gain, satisfaction and hedonism are consistent with our hypothesis, since these factors were better evaluated when the verbal communication has occurred. The user experience remains, due to its very subjective nature and its dependency to each specific context, a notion that is quite difficult to assess. Nevertheless, our current findings will guide our future work both in terms of new measurements and types of task to be performed by the participants.

Interestingly, we have met some unforeseen difficulties during this experiment. For example, the use of Occulus Rift has caused strong motion sickness to some of our participants. This has resulted in the loss of 8 data sets because these participants could not remain immersed along the 10 minutes required for the validation of their answers to questionnaires. We also planned to collect physiological and neurological responses of our participants by using electroencephalogram, electrocardiogram and galvanic skin response sensors. Unfortunately, some technical issues, such as sensors' signal interferences and unreliable sticky patches, made this attempt unsuccessful. These difficulties explain the sole use of survey questionnaires and quantitative data in this study.

The main objective of our future work is to get a deeper understanding of the impact of verbal and non-verbal communication on social interaction and on the quality of experience. Besides increasing QoE and understanding impacting factors as explained in the introduction, we intent to improve the design of new applications involving social interactions. For example: interactions between patient and psycho-sociologist along the therapeutic treatment inside 3D-IVEs. Another option would consist to make a second attempt to implement a range of objective measurements, such as above-mentioned physiological and neurological responses, that were not evaluated during this study. Beside these objective measurements, we would add qualitative data through complementary focus groups and semi-structured interviews. Furthermore, we plan to expose participants to a visual feedback of the physio-emotional state of their pairs through a colored artifact representing one's aura. Such artifact development would propagate an interpretation of the physio-emotional state of one and other and would favor a behavioral reaction like in the famous translucence study [27].

Acknowledgments

This empirical study has been carried out in the context of a thesis grant from Laval local authorities, France. Authors wish to acknowledge their gratitude and appreciation to Laval local authorities and the Ecole de Design Nantes Atlantique (Nantes, France) for their active support. Authors wish to particularly thank Dr. Lionel Bennes (OPTIS Imagine, Antibes, France) for his help in the experimentation process.

References

[1] N. Yee, N. Ducheneaut, L. Nelson, and P. Likarish, "Introverted elves & conscientious gnomes," presented at the the 2011 annual conference, New York, New York, USA, 2011, p. 753.

[2] A. Guye-Vuill me, T. K. Capin, I. S. Pandzic, N. M. Thalmann, and D. Thalmann, "Nonverbal communication interface for collaborative virtual environments," *Virtual Reality*, vol. 4, no. 1, pp. 49–59, Mar. 1999.

[3] N. Ducheneaut, N. Yee, E. Nickell, and R. J. Moore, "Alone together?: exploring the social dynamics of massively multiplayer online games," presented at the The SIGCHI Conference on Human Factors in Computing Systems 2006, New York, New York, 2006, pp. 407–416.

[4] B. Rubio Ballester, S. Bermúdez i Badia, and P. F. M. J. Verschure, "Including Social Interaction in Stroke VR-Based Motor Rehabilitation Enhances Performance: A Pilot Study," *Presence: Teleoperators and Virtual Environments*, vol. 21, no. 4, pp. 490–501, Nov. 2012.

[5] M. Argyle, *Bodily communication*, 2nd ed. Routledge, 1988.

[6] J. N. Bailenson, N. Yee, D. Merget, and S. Ralph, "The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interactions," *Presence: Teleoperators and Virtual Environments*, vol. 15, no. 4, pp. 359–372, Sep. 2006.

[7] D. L. Scapin, B. Senach, B. Trousse, and M. Pallot, "User Experience: Buzzword or New Paradigm?," presented at the The Fifth International Conference on Advances in Computer-Human Interactions 2012, Valencia, Spain, 2012.

[8] W. Wu, A. Arefin, R. Rivas, K. Nahrstedt, R. Sheppard, and Z. Yang, "Quality of experience in distributed interactive multimedia environments: toward a theoretical framework," presented at the ACM Multimedia Conference 2009, New York, New York, USA, 2009, pp. 481–490.

[9] M. Garau, "The impact of avatar fidelity on social interaction in virtual environments," University of London, London, United Kingdom, 2003.

[10] M. Pallot, B. Poussard, O. Christmann, and S. Richir, "Augmented Sport: Exploring Collective User Experience," presented at the Virtual Reality International Conference 2013, 2013.

[11] M. Pallot, K. Pawar, and R. Santoro, "A User Experience Framework and Model within Experiential Living Labs for Internet of Things," presented at the 2013 19th International ICE Conference on Engineering, Technology and Innovation (ICE), 2013.

[12] R. M. Baños, C. Botella, and M. Alcañiz, "Immersion and emotion: their impact on the sense of presence," *CyberPsychology & Behavior*, vol. 7, no. 6, pp. 734–741, Feb. 2005.

[13] M. E. Hoque, M. Courgeon, J.-C. Martin, and B. Mutlu, "MACH: My Automated Conversation coacH," *web.media.mit.edu*, Jun-2013. [Online]. Available: http://web.media.mit.edu/~mehoque/Publications/13.Hoque-etal-MACH-UbiComp.pdf. [Accessed: 21-Jun-2013].

[14] D.-P. Pertaub, M. Slater, and C. Barker, "An Experiment on Public Speaking Anxiety in Response to Three Different Types of Virtual Audience," *Presence: Teleoperators and Virtual Environments*, vol. 11, no. 1, pp. 68–78, Feb. 2002.

[15] R. E. Guadagno, J. Blascovich, and J. N. Bailenson, "Virtual humans and persuasion: The effects of agency and behavioral realism," *MP*, vol. 10, no. 1, 2007.

[16] J. N. Bailenson, A. C. Beall, and J. Blascovich, "Gaze and task performance in shared virtual environments," *the journal of Visualization and Computeur Animation*, vol. 13, no. 5, pp. 313–320, Dec. 2002.

[17] J. N. Bailenson, J. Blascovich, and A. C. Beall, "Interpersonal distance in immersive virtual environments," *PSPB*, vol. 29, no. 7, pp. 819–833, Jul. 2003.

[18] E. L.-C. Law, V. Roto, M. Hassenzahl, A. P. O. S. Vermeeren, and J. Kort, “Understanding, scoping and defining user experience,” presented at the The SIGCHI Conference on Human Factors in Computing Systems 2009, New York, New York, USA, 2009, pp. 719–728.
[19] A. Kankainen, *Thinking model and tools for understanding user experience related to information appliance product concepts*, vol. 118. Alto University Library, 2002.
[20] A. Gaggioli, M. Bassi, and A. Delle Fave, “Quality of experience in virtual environments,” Ios Press, Being There: Concepts, effects and measurement of user presence in synthetic environments, 2003, pp. 122–132.
[21] K. Nahrstedt, A. Arefin, R. Rivas, P. Agarwal, Z. Huang, W. Wu, and Z. Yang, “QoS and resource management in distributed interactive multimedia environments,” *MTA*, vol. 51, no. 1, pp. 99–132, Jan. 2011.
[22] P. Sweetser and P. Wyeth, “GameFlow: a model for evaluating player enjoyment in games,” *CIE*, vol. 3, no. 3, pp. 3–3, Jul. 2005.
[23] M. Lombard and T. Ditton, “At the heart of it all: The concept of presence,” *JC-MC*, vol. 3, no. 2, p. 0, Sep. 1997.
[24] C. Heeter, “Being there: the subjective experience of presence,” *Presence: Teleoperators and Virtual Environments*, vol. 1, no. 2, pp. 262–271, May 1992.
[25] T. Schubert, F. Friedmann, and H. Regenbrecht, “The Experience of Presence: Factor Analytic Insights,” *Presence: Teleoperators and Virtual Environments*, vol. 10, no. 3, pp. 266–281, Jun. 2001.
[26] J. Freeman, J. Lessiter, E. Keogh, and J. Davidoff, “A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory,” *Presence: Teleoperators and Virtual Environments*, vol. 10, no. 3, pp. 282–297, Jun. 2001.
[27] T. Erickson and W. A. Kellogg, “Social translucence: an approach to designing systems that support social processes,” *TOCHI*, vol. 7, no. 1, pp. 59–83, Mar. 2000.
[28] F. Biocca, C. Harms, and J. K. Burgoon, “Towards A More Robust Theory and Measure of Social Presence: Review and Suggested Criteria,” *Presence: Teleoperators and Virtual Environments*, vol. 12, no. 5, pp. 456–480, Oct. 2003.
[29] A. M. Grinberg, J. S. Careaga, and M. R. Mehl, “Social engagement and user immersion in a socially based virtual world,” *CHB*, vol. 36, pp. 479–486, Jul. 2014.
[30] F. Biocca and C. Harms, “Defining and measuring social presence: Contribution to the networked minds theory and measure.,” presented at the 5th International Workshop on Presence 2002, 2002, pp. 7–36.
[31] P. Siriaraya and C. Siang Ang, “Age differences in the perception of social presence in the use of 3D virtual world for social interaction,” *IwC*, vol. 24, no. 4, pp. 280–291, 2012.
[32] T. Schubert and F. Friedmann, “Embodied presence in virtual environments,” R. Paton and I. Nielsen, Eds. Visual representations and interpretations, 1999, pp. 268–278.
[33] T. Schubert, F. Friedmann, and H. Regenbrecht, “Decomposing the sense of presence: Factor analytic insights,” presented at the 2nd International Workshop on Presence 1999, 1999.
[34] F. Biocca, C. Harms, and J. Gregg, “The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity,” presented at the 4th International Workshop on Presence 2001, 2001.
[35] B. J. Winer, D. R. Brown, and K. M. Michels, *Statistical Principles in Experimental Design*, 3rd ed. New York: McGraw Hill, 1991.
[36] B. Rey, V. Parkhutik, J. Tembl, and M. Alcañiz, “Breaks in presence in virtual environments: An analysis of blood flow velocity responses,” *Presence: Teleoperators and Virtual Environments*, vol. 20, no. 3, pp. 273–286, Jun. 2011.
[37] T. Schubert, H. Regenbrecht, and F. Friedmann, “Real and illusory interactions enhance presence in virtual environments,” *Presence: Teleoperators and Virtual Environments*, vol. 11, no. 4, pp. 425–434, Aug. 2002.