Training doctors' social skills to break bad news: evaluation of the impact of virtual environment displays on the sense of presence

Magalie Ochs, Daniel Mestre, Grégoire de Montcheuil, Jean-Marie Pergandi, Jorane Saubesty, Evelyne Lombardo, Daniel Francon, Philippe Blache

▶ **To cite this version:**

Magalie Ochs, Daniel Mestre, Grégoire de Montcheuil, Jean-Marie Pergandi, Jorane Saubesty, et al.. Training doctors' social skills to break bad news: evaluation of the impact of virtual environment displays on the sense of presence. Journal on Multimodal User Interfaces, 2019, 13 (1), pp.41-51. 10.1007/s12193-018-0289-8 . hal-02185446

HAL Id: hal-02185446
https://hal.science/hal-02185446

Submitted on 21 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Noname manuscript No.
(will be inserted by the editor)

Training doctors' social skills to break bad news: Evaluation of the impact of virtual environment displays on the sense of presence

Magalie Ochs · Daniel Mestre · Grégoire de Montcheuil · Jean-Marie Pergandi · Jorane Saubesty · Evelyne Lombardo · Daniel Francon · Philippe Blache

the date of receipt and acceptance should be inserted later

Abstract . The way doctors deliver bad news has a significant impact on the therapeutic process. In order to facilitate doctor's training, we have developed an embodied conversational agent simulating a patient to train doctors to break bad news. In this article, we present an evaluation of the virtual reality training platform comparing the users' experience depending on the virtual environment displays: a PC desktop, a virtual reality headset, and four wall fully immersive systems. The results of the experience, including both real doctors and naive participants, reveal a significant impact of the environment display on the perception of the user (sense of presence, sense of co-presence, perception of the believability of the virtual patient), showing, moreover, the different perceptions of the participants depending on their level of expertise.

Keywords Virtual reality · Presence · Training · Embodied Conversational Agent · Health domain · Breaking bad news

1 Introduction

Many works have shown that doctors should be trained not only to perform medical or surgical acts but also to develop skills in communication with patients [**?**,**?**,**?**]. Among all possible bad new, doctors can be faced with the complex situation of announcing a damage associated to a care that can engage their responsibility: unforeseeable medical situation, dysfunction, medical error, etc. The way *doctors deliver bad news related to damage associated with care* has a significant impact on the therapeutic process: disease evolution,

M. Ochs, D. Mestre, G. de Montcheuil, J-M Pergandi, J. Saubesty, E. Lombardo, P Blache
Aix Marseille Université, Université de Toulon, CNRS, LIS UMR 7020, LPL UMR 7309, ISM UMR7287

D. Francon
Département anesthésie réanimation, Institut Paoli Calmettes, Marseille, France

adherence with treatment recommendations, litigation possibilities [**?**]. However, both experienced clinicians and medical students consider this task as difficult, daunting, and stressful.

Training health care professional to break bad news is now recommended by several national agencies (e.g. the French National Authority for Health, HAS)[1]. Such trainings are organized as workshops during which doctors disclose bad news to actors playing the role of patients. This solution is complex to implement: it requires several persons, it is costly, and time consuming (each 30 mn. session requires one hour of preparation). Our project[2] aims at developing *a virtual reality training system* with an embodied conversational agent playing the role of a virtual patient.

In this paper, we present an evaluation of a first *semi-autonomous* virtual reality training system inhabited by a virtual patient and developed to give the capabilities to doctors to simulate breaking bad news situation. The *semi-autonomous* system includes both automatic and manual modules, making it possible to simulate a fully automatized human-machine interaction. Implemented on three different virtual environment displays (PC, virtual reality headset, and an immersive virtual reality room), the doctors can interact in natural language with a virtual patient that communicates through verbal and non-verbal behavior. A first evaluation has been conducted to evaluate the capacity of the training system to offer an immersive experience to users in this specific task of breaking bad news. The type of participants (naive users and real doctors) enabled us to evaluate the impact of the participant's expertise on their perception of the interaction. Considering the three different virtual environment displays with different degrees of immersion, we evaluated the effects of the display on the user's *sense of presence*.

The paper is organized as follows. In the next section, we present related works on virtual patients used for doctor's training. In Section **??**, we discuss the theoretical background on the evaluation of the sense of presence. Section **??** is dedicated to the presentation of the virtual training system. The evaluation of the system and the results are presented Section **??**. We conclude in Section **??**.

2 Related works: Virtual patients to train doctors' social skills

Several studies have shown that embodied conversational agents are perceived as social entities leading users to show behaviors that would be expected in human-human interactions [**?**,**?**]. Several virtual agents embodying the role of virtual patients have already been proposed for use in clinical assessments, interviewing and diagnosis training [**?**,**?**,**?**]. Indeed, previous works have shown that doctors demonstrate non-verbal behaviors and respond empathetically to

[1] The French National Authority for Health is an independent public scientific authority with an overall mission of contributing to the regulation of the healthcare system by improving health quality and efficiency.

[2] ACORFORMed, http://www.lpl-aix.fr/˜ acorformed/.

a virtual patient [?]. In this domain, the research has mainly focused on the anatomical and physiological models of the virtual patient to simulate the effects of medical interventions or on models to simulate particular disorder (e.g. [?,?] or the eViP European project[3]). In our project, we focus on a virtual patient to train doctors to deliver bad news.

Recent works [?,?] showed that virtual agents could help human beings improve their social skills. For instance, in [?], a virtual agent is used to train kids to adapt their language register to the situation. In the European project TARDIS [?], an ECA endowed the role of a virtual recruiter is used to train young adults to job interview. More specifically in the context of training doctors to break bad news, a first study [?] has analyzed the benefits of using a virtual patient to train doctors to break a diagnosis of breast cancer. The results show significant improvements of the self-efficacy of the medical trainees. The main limit of the proposed system, highlighted by the participants, is the lack of non-verbal behaviors of the patients simulated in the limited Second Life environment. In this project, our objective is, in particular, to simulate the non-verbal expressions of the virtual patient to improve the believability of the virtual character and the immersive experience of the doctor.

Most of the embodied conversational agents used for health applications have been integrated in 3D virtual environments on PCs. Virtual reality in health domain is particularly used for virtual reality exposure therapy (VRET) in the treatment for anxiety and specific phobias (e.g. [?]) but also for social patient perspective taking [?]. In our project, in order to offer an immersive experience to the doctor, we have integrated the virtual patient in a virtual reality environment.

Some research works have compared the experience of the users in virtual environments, depending on the display used. For instance, in the case of exposure therapy (typically acrophobia), an immersive virtual reality room has appeared as more effective than a virtual reality headset [?,?]. In the context of education, in comparison with a PC, the virtual reality room seems to lead to better learning [?]. In [?] and [?], the authors have shown that, for a specific navigation task, the PC is more appropriate that the virtual reality headset that can generate cybersickness. In a recent work presented in [?], the authors have compared PC and virtual reality headset in the context of a serious game. The results did not reveal significant differences in terms of learning but showed a stronger engagement and immersion with the headset than with a PC.

Finally, as far as we know, the comparison of users' experiences interacting with a virtual patient in the context of social competencies training, with different virtual environment displays, has not been analyzed. In this paper, we present such an evaluation in the context of breaking bad news. To evaluate the user's experience, we focus on the sense of presence. In the next section, we present this concept and the existing tools to measure it.

[3] http://virtualpatients.eu/

3 Theoretical background: the sense of presence

3.1 Definition of the sense of presence

In the literature, *technological* and *physical* immersion [**?**] are distinguished, based on the characteristics of the device and caused particularly by 360 degrees displays, and psychological immersion [**?**] which is, to some degree, independent of the device (a book, projecting us in a virtual world, can trigger a psychological immersion, without technological and physical immersion). This type of immersion is called *sense of presence* and approaches the concept of *flow* [**?**] that makes the user lose track of time and space.

Two different schools of thought can be distinguished concerning the definition of immersion. First, in [**?**], the authors consider immersion as a psychological state, as the perception of *being in, to be surrounded by*. Immersion includes for these authors the insulation from the physical environment, the perception of a feeling of being *included* in the virtual environment, the natural state of the interactions, a perception of control, and the perception of movement in a virtual environment. Another approach considers immersion in a technological view: immersion would strongly be linked to technology [**?**,**?**,**?**]. In this study, our definition to the sense of presence relies on that given in [**?**].

Several factors are identified as affecting the sense of presence: (1) *the ease of interaction*: interaction correlates with the sense of presence felt in the virtual environment [**?**]; (2) *the user control*: the sense of presence increases with the sense of control [**?**]; (3) *the realism of the image*: the more realistic virtual environment is, the stronger is the sense of presence [**?**]; (4) the *duration of exposure*: prolonged exposure beyond 15 minutes with the virtual environment does not give the best result for the sense of presence with a HMD (Head Mounted Display) and there is even a negative correlation between the prolonged exposure in the virtual environment and the sense of presence [**?**]; (5) the *social presence and social presence factors*: the social presence of other individuals (real or avatars), and the ability to interact with these individuals increases the sense of presence [**?**]; (6) the *the quality of the virtual environment*: quality, realism, the ability of the environment to be fluid, to create interaction are key factors for the sense of presence of the user [**?**]. Two other factors are more particularly related to the individual perception, and contextual and psychological factors that should be taken into account during the evaluation of presence [**?**]. In the next section, we introduce the different questionnaires available to measure these factors.

3.2 Presence Questionnaires

To test the sense of presence, several questionnaires have been proposed. Four of them are canonical since they have been tested several times in other research and are statistically significant: the canonical presence test of Witmer and Singer [**?**], the ITC-SOPI canonical test [**?**] that evaluates the psycho-

logical immersion, the Slater-Usoh-Steed (SUS) questionnaire to evaluate the spatial presence, and the canonical test IGroup Presence Questionnaire (IPQ) [**?**]. The latest has been used in our study to evaluate our training system. This test aims at evaluating three variables dependent on presence factors: the spatial presence, the involvement in the device, and the realism of the device. The test is composed of 14 questions, some of them are taken directly from the Presence Questionnaire [**?**] and the SUS questionnaire [**?**]. In the last version, another variable, evaluating global presence was added in the test. This test has the advantage to contain few questions (only 14) while including the main presence factors of the other canonical tests.

However, one limit of the IPQ test is the lack of the evaluation of the notion of *copresence*. Copresence, also commonly called *social presence*, can be defined as "the sense of being and acting with others in a virtual space" [**?**][4]. In our context, we are interested in evaluating the sense of copresence of the participants with the virtual patient depending on the used virtual environment display (PC, virtual reality headset, or virtual reality room). In order to evaluate copresence, we have used the test proposed in [**?**], that measures social presence through the following variables: *perceived copresence*, *embarrassment* to measure the social influence of the agent, and *likability* of the virtual representation of the agent. In [**?**], the authors have shown that this self-report questionnaire is effective "to measure how people perceive an embodied agent".

As highlighted in [**?**], immersion and more particularly the sense of presence reflecting the involvement of the users contribute to positive learning outcomes. In our context the sense of presence and co-presence is all the more important given that we aim at simulating a "real" communication situation in a virtual environment inhabited by a virtual patient.

In the next section, we present in more details the developed virtual reality training platform.

4 A Semi-Autonomous Virtual Reality Training Platform

As concerns the general architecture of the training platform, we have developed a *semi-autonomous* platform. The architecture is described in Figure **??**. The platform is *semi-autonomous* because some modules of the system are automatic (for example the dialogue generation) while others are manual (controlled by a trained operator). In particular, the speech recognition and the comprehension modules are simulated by a human: the doctor verbal production is interpreted in real time by the operator who selects the adequate input signal to be transmitted to the dialogue system. Indeed, these modules may be particularly critical in case of failure of an automated module and strongly damage the interaction. They represent moreover a particularly difficult part of the system(currently underdevelopment). Replacing these modules by the

[4] Note that no consensus exists on the notion of co-presence. A detailed discussion on the different definitions can be found in [**?**]

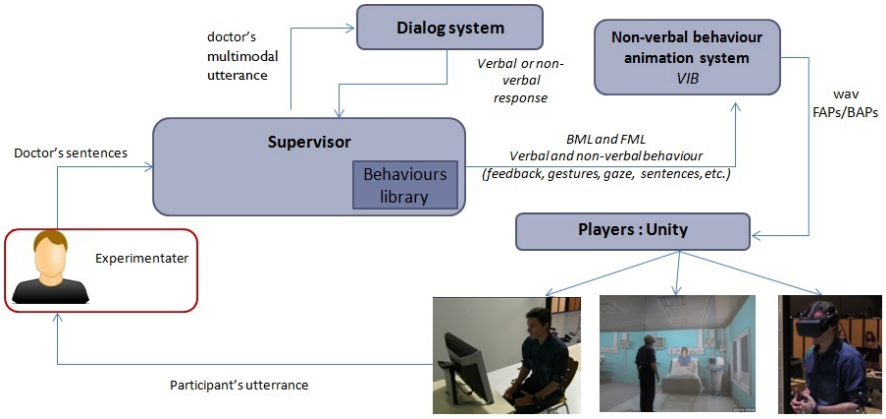

Fig. 1: Overall architecture of training platform. In the figure, the abbreviations BML and FML correspond to XML-based language of the SAIBA framework, a standard in the ECA research community, used to describe the semantic information of the verbal and non-verbal behavior (in terms of communicative intentions) for the FML (Function Markup Language) and in terms of verbal and non-verbal signals for the BML (Behavior Markup Language) [**?**]. The abbreviations FAP and BAP correspond respectively to Face and Body Animation Parameters in the MPEG-4 standard.

operator comes to a perfect speech recognition and comprehension. This makes it possible to completely control the corresponding parameters and concentrate on the evaluation of the others modules, such as the dialog supervision and the non-verbal behavior of the virtual patient. Moreover, it renders possible the evaluation of the overall interaction (e.g. presence and copresence).

A specific interface has been designed for this purpose to enable the experimenter to select the sentences semantically matching what has been said by the doctors (Figure **??**). The interface contains 136 prototypical sentences (or patterns) organized into different dialog phases: greetings, asking the patient's feelings, description of the surgical problem, description of the remediation. These sentences have been defined based on the analysis of a transcribed corpus of doctor-patient interaction (the corpus is described in details in [**?**]). Each prototypical sentence encodes a family of possible utterances, as identified in the corpus. The sentences are encoded into an XML file. Keyboard shortcuts are associated to each sentence/pattern, and can be configured in order to be easily selected by the experimenter. Several pre-tests have been built to test the interface and train the experimenter. Note that at the difference with a "Wizard of Oz" setup, the experimenter does not select the virtual patient's reaction but only send to the dialog model the recognized doctor's sentence.

The dialogue system then generates a sequence of instructions, to be sent to a non-verbal behavior animation system called VIB [**?**]. This system computes

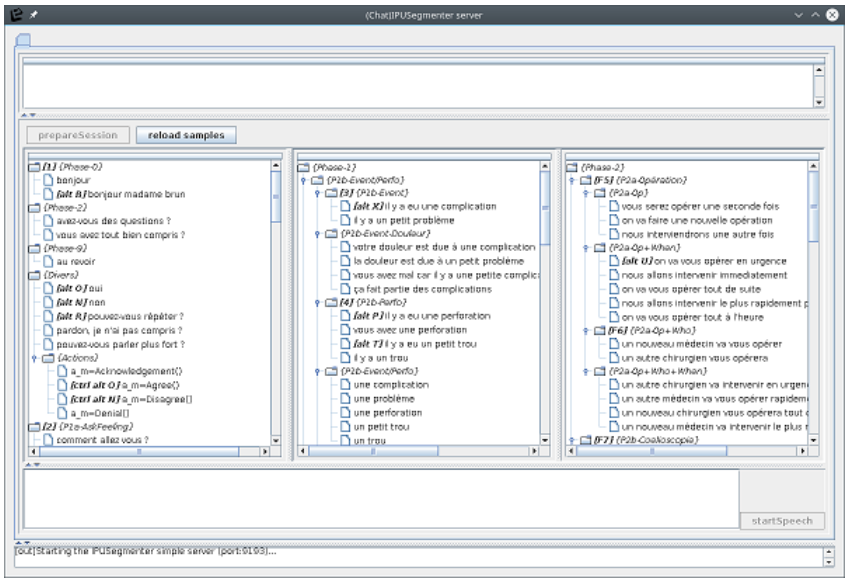

Fig. 2: Screen-shot of the interface of the experimenter to select the corresponding doctor's recognized sentences

the animation parameters (*Facial Animation Parameters* - FAP - and *Behavioral Animation Parameters* - BAP) to animate the face and the body of the virtual patient. The result is encoded in XML and describes the communicative intention to perform (encoded in FML, *Function Markup Language*) as well as the non-verbal signals to express (encoded in BML, *Behavior Markup Language*). Moreover, the VIB system contains a text-to-speech synthesis [**?**] for generating the speech, in synchronization with the non-verbal behavior (including lips animation).

In order to experiment as broadly as possible the validity of the approach, we have implemented the virtual patient on different virtual environment displays: PC, virtual reality headset (an Oculus Rift), and a virtual reality room. The virtual reality room, a CAVE, is constituted of a 3m deep, 3m wide, and 4m high cubic space with three vertical screens and a horizontal screen (floor). A cluster of graphics machine makes it possible to deliver stereoscopic, wide-field, real-time rendering of 3D environment, including spatial sound. This offers an optimal sensorial immersion of the user. The environment has been designed to simulate a real recovery room where the breaking bad news are generally performed. The virtual agent based on the VIB platform has been integrated in by means of the *Unity* player. In the next section, we present the evaluation of the training platform.

5 Evaluation of the Semi-Autonomous Training Platform

5.1 Method

Participants. In total, 22 persons (16 males, 6 females) with a mean age of 29 years (SD:10.5) participated in the experimentation. Some participants (12) have been recruited in Aix-Marseille University. Ten of them (7 males, 3 females) were real doctors recruited in a medical institution. These participants had already an experience in breaking bad news with real patients. The participants were not paid. They were recruited on a voluntary basis and signed an informed consent form.

Design. The design of the experiment consisted in *one independent variable* corresponding to the virtual environment display used for the interaction that could be either a PC, a virtual reality headset or a virtual reality room (as illustrated in Figure **??**). Note that for each condition, the participant was positioned in the same space (in the virtual reality room that was switched off in the PC and headset condition). In the PC condition, the participants were sat in a chair. In a *within-subject* configuration, each participant interacted

Fig. 3: Participants interacting with the virtual patient with different virtual environment displays (from left to right): virtual reality headset, virtual reality room, and PC.

with each virtual environment display. We counterbalanced the order of the use of each display in order to avoid an effect of the display condition order on the results.

Equipment. The virtual room in which participants were immersed is illustrated in Figure **??**. The virtual room displayed in the PC condition is illustrated Figure **??**. The same embodied conversational agent was used for each condition. The participants were filmed using a video camera. Their gestures and head movements were digitally recorded from the tracking data: their head, elbows and wrists were equipped with tracked targets using the cave real-time tracking system. A high-end microphone synchronously recorded the participant's verbal expression. The participant's behavior was recorded in order to be able to replay the interaction with a custom 3D video playback

Fig. 4: The virtual room with the embodied conversational agent for the headset and virtual reality room condition.

Fig. 5: Virtual room with the embodied conversational agent displayed in the PC condition.

player (Figure **??**) as described below. As for the virtual agent, its gesture and verbal expressions were recorded from the Unity Player. The visualization of the interaction can be done through a 3D video playback player we have developed (Figure **??**). This player replays synchronously the animation and verbal expression of the virtual agent as well as the movements and video of the participant.

Procedure. When participants arrived at the laboratory, an experimenter sat them down and presented them the instructions. Each participant interacted

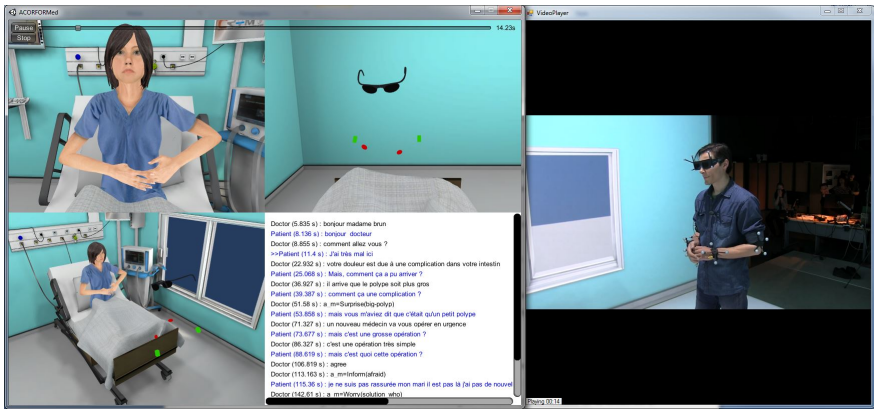

Fig. 6: 3D video playback player

with the systems 3 times with three different displays: PC, virtual reality headset, and virtual reality room. The task of the participants was to announce a digestive perforation after a gastroenterologic endoscopy in the immediate post operative period[5]. Before the interaction, written instructions were presented to the participants: the role they have to play is a doctor that had just operated the virtual patient to remove a polyp in the bowel. A digestive perforation occurred during the endoscopy. These written instructions explained precisely the causes of the problem, the effects (pain), and the proposed remediation (a new urgent surgery). Participants were asked to read the instructions several times as well as before each interaction. The understanding was verified by means of an oral questionnaire. Each participant had the instruction to announce this medical situation to the virtual patient three times with the three different displays. The duration of each interaction was in average 3mn16s (an example of interaction is presented on the ACORFORMed site).

Measures. In order to evaluate the participant's experience, we asked (after each experimental session) the participants to respond to different questions on their subjective experience to measure their sense of *presence* (with the *IGroup Presence Questionnaire*, IPQ [**?**], described in Section **??**) and their sense of *copresence* [**?**] (described in Section **??**). These questionnaires were completed using questions extracted from the questionnaire proposed in [**?**]. These questions enabled us to measure the perception of the believability of the virtual patient and the perception of the communication. Moreover, as proposed in [**?**], we measured the user's perception if her/his performance in delivering bad news by adapting the questions proposed in [**?**] to our context (e.g. "I had difficulty to deliver bad news to the virtual patient"). In total, the

[5] The scenario has been carefully chosen with the medical partners of the project for several reasons (e.g. the panel of resulting damages, the difficulty of the announcement, its standard characteristics of announce).

participants responded to 31 questions after each interaction through Likert scale of 5 points.

5.2 Results

In this section, we present the main significant results arising from the post-experience questions presented to the participants. First, we analyzed the three categories of the IGroup Presence Questionnaire (IPQ).

The spatial presence. The spatial presence characterizes the "sense of being there" [**?**]. Concerning this spatial dimension, we found a significant effect of the experimental setup (Figure **??**). Post-Hoc analyses showed that the PC-screen gave significantly inferior scores, as compared to CAVE (virtual reality room) and HMD setups ($p < .0001$). However, the difference between CAVE and HMD did not reach statistical significance ($p = .06$).

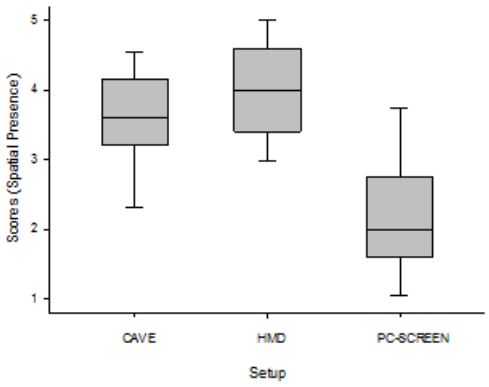

Fig. 7: Boxplot depicting the Spatial Presence scores as a function of the setup used

Involvement. Involvement characterizes the attention to real and virtual environment (e.g. "I was totally captivated by the virtual world") [**?**]. Concerning involvement, we also observed a significant effect of display conditions ($p < .0001$). Involvement is lower for the PC setup. Moreover, there is also a main effect of group: experts (doctors) are overall more involved than naive participants ($p < .02$). Post-Hoc analysis (Bonferroni) shows that experts are significantly more involved than naive participants for the CAVE setup (Figure **??**).

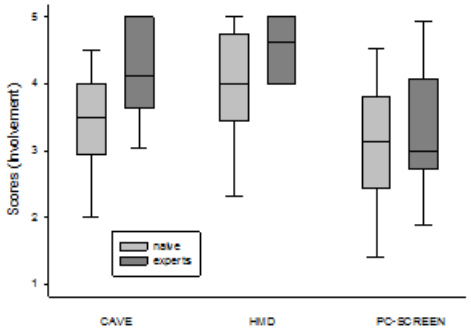

Fig. 8: Boxplot depicting the Involvement scores as a function of the setup used.

Realness. The realness factor refers to a comparison between the virtual and the real world (e.g. "How real did the virtual world seem to you?") [**?**]. For the realism score, we observed a pattern of results similar to the previous ones, realism being judged as reduced in the PC setup, compared to HMD and CAVE setups, these latter being not different from each other.

Co-presence. Co-presence is measured through 3 factors: perception of co-presence, embarrassment of the user and likability of the virtual character (Section **??**). As concerns *perceived co-presence*, again we found a pattern of results similar to spatial presence or realism (Figure **??**). For the *embarrass-*

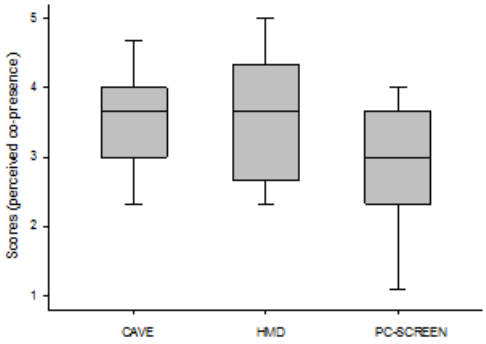

Fig. 9: Boxplot depicting the co-presence scores as a function of the setup used.

ment and *likability* scores, we found no significant difference between setups

or group. Moreover, average scores were overall low, with average values of 2.5 ($SD = 1.4$) and 3.03 ($SD = 1$) for *embarrassment* and *likability*, respectively (on a Likert scale of 5 points). These values mean that participants were not specially embarrassed in front of the virtual patient.

Perception of the believability of the virtual patient and of the communication. To the question "was the virtual patient credible as compared to real patient", scores are statistically higher in the CAVE setup, as compared to the PC setup, HMD and PC being not different. Moreover, there is a trend ($p = .07$) for experts to give higher scores than naive participants, especially in the CAVE setup (Figure **??**). To the question "was the virtual patient reactive to what

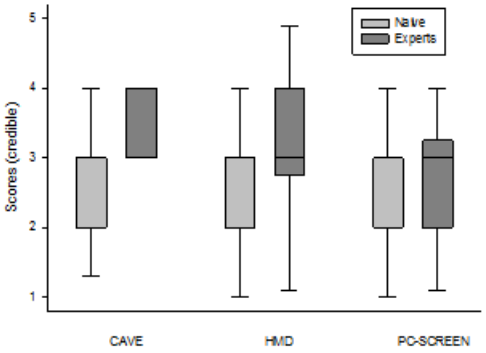

Fig. 10: Boxplot depicting the perception of the believability of the virtual patient scores as a function of the setup used.

you said", the participants responded with an average score around 3 (on a Likert scale of 5 points), meaning that they judged that the virtual patient was moderately reactive. This can easily be explained by the characteristics of the platform since the virtual patient does not currently express feedback (e.g. head nodes) during the participant's speech, an important element for the flow of the communication. To the question"do you think you and the virtual patient understood each other", there was no differences between setups and groups. Average scores are equal to 2.75 ($SD = 1$), suggesting that mutual comprehension was not completely satisfactory, again possibly due to the lack of virtual patient listener's behavior.

Perception of the user's own performance . Finally, as concerns the evaluation by the participants of their own performance while interacting with the virtual patient, to the question "how well do you think you did explain the problem to the virtual patient", overall and with no significant difference between groups and setups, the participants gave average scores of 3.34 ($SD = .93$). Coherently, to the question "I had difficulties delivering the bad news to the virtual

patient", again, there was no significant difference between groups and setups. Average score are equal to 2.33 ($SD = 1$). This pattern of results means that they were rather satisfied with their performance.

Effect of the repeated experience. In order to assess the effect of the repeated experience on the participants, we computed the conversation time of the participants for each session (first second and third). The results are illustrated Table 1.

Table 1: Conversation time as a function of task repetition

Participants	Session 1	Session 2	Session 3	Sign
Naive	190.03 (62.89)	175.97 (45.81)	166.05 (41.78)	*P=0.019*
Experts	243.07 (46.72)	229.31 (42.69)	222.19 (46.69)	P=0.33

Interestingly, the results show significant differences between the experts and the naive participants. There is no significant difference for the experts, however the naive participants seem to speak less during the third session compared to the second and the first ones. These results may suggest that the naive may get bored by the repeated experience but not the experts seem to remain involved over repeated sessions.

5.3 Discussion

Overall, participants give higher scores of presence and co-presence for HMD and CAVE, as compared to the PC setup. This is coherent with the general idea that these setups are more *immersive*, thus leading to higher presence: the spatial presence, the involvement, the realness but also the co-presence. HMD and CAVE are not statistically different. However, the group factor (naive vs expert participants) shows interesting effects. First, experts tend to be more involved than naive participants, especially for the CAVE setup. One suggestion is that the CAVE enables experts to be immersed in a "familiar" environment, without being isolated from the real environment. The fact that the experts also judge the virtual patient to be more credible in the CAVE setup argues in the same direction. Experts judge the virtual patient as credible, with respect to their motivation to engage in the conversation, more so in the CAVE setup.

In a nutshell, the virtual reality room and the virtual reality headset appear as the most appropriate virtual reality environment displays for the training of doctors to break bad news. In particular, the virtual reality room seems to enhance the doctor's experience (sense of presence and perception of the virtual patient) compared to the virtual reality headset and the PC. The results reveal a significant impact of the expertise of the participants, showing the potential effect of "familiar" context on the virtual reality experience. Considering the number of participants, further and more extended evaluation of an entirely autonomous system will enable us to confirm the results of this experiment

6 Conclusion

In conclusion, we presented in this article a semi-autonomous system to train doctors to break bad news with a virtual patient. The evaluation of such a system by considering different virtual reality environment displays, enabled us to identify the most appropriate display for this task in term of users' experience. The results show that the virtual reality room is particularly suitable for doctors compared to naive participants.

We are currently developing a fully autonomous training platform (in particular the comprehension and generation modules). We have already used the corpus obtained during the experimentation to train and test the speech recognition system, in order to ensure that the speech recognition system can accurately recognize the participants. We also verified that the recognized words and sentences activate correctly the expected rules in the dialog model. Moreover, the corpus of the experimentation - and in particular the recordings of participants' speech and head and body movements - is currently being analyzed to compare the participants' verbal and non-verbal behavior through the different devices and to try to link objective measures (like the movements) to the subjective measure of presence and co-presence. In order to improve the communication, we are integrating a feedback model in the virtual patient to give it the capabilities to express backchannels during the doctor's speech [**?**].

The final step is the evaluation of the fully autonomous training platform and of the trainee's learning. The trainee evaluation is an entire research subject. An evaluation grid will be defined, starting from the existing one currently used in the hospitals: the "Affective Competency Score" [**?**]. The ACS will be scored by the trainees to measure their self-efficacy before and after a session with the virtual patient. Professional observers will also rate the ACS to evaluate the trainees' performances.

Acknowledgements This work has been funded by the French National Research Agency project ACORFORMED (ANR-14-CE24-0034-02) and supported by grants ANR-16-CONV-0002 (ILCB), ANR-11-LABX-0036 (BLRI) and ANR-11-IDEX-0001-02 (A*MIDEX).

References

1. Anderson, K., André, E., Baur, T., Bernardini, S., Chollet, M., Chryssafidou, E., Damian, I., Ennis, C., Egges, A., Gebhard, P., et al.: The tardis framework: intelligent virtual agents for social coaching in job interviews. In: Advances in computer entertainment, pp. 476–491. Springer (2013)
2. Andrade, A., Bagri, A., Zaw, K., Roos, B., Ruiz, J.: Avatar-mediated training in the delivery of bad news in a virtual world. Journal of palliative medicine **13**(12), 1415–1419 (2010)
3. Aylett, M.P., Pidcock, C.J.: The cerevoice characterful speech synthesiser sdk. In: IVA, pp. 413–414 (2007)
4. Baile, W., Buckman, R., Lenzi, R., Glober, G., Beale, E., Kudelka, A.: Spikes-a six-step protocol for delivering bad news: application to the patient with cancer. Oncologist **5**(4), 302–311 (2000)

5. Bailenson, J.N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., Blascovich, J.: The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. Presence: Teleoperators and Virtual Environments **14**(4), 379–393 (2005)
6. Billinghurst, M., Weghorst, S.: The use of sketch maps to measure cognitive maps of virtual environments. In: Virtual Reality Annual International Symposium, 1995. Proceedings., pp. 40–47. IEEE (1995)
7. Buttussi, F., Chittaro, L.: Effects of different types of virtual reality display on presence and learning in a safety training scenario. IEEE transactions on visualization and computer graphics (2017)
8. Bystrom, K.E., Barfield, W., Hendrix, C.: A conceptual model of the sense of presence in virtual environments. Presence: Teleoperators and Virtual Environments **8**(2), 241–244 (1999)
9. Cadoz, C.: Les réalités virtuelles. Flammarion (1994)
10. Csikszentmihalyi, M.: Toward a psychology of optimal experience. In: Flow and the foundations of positive psychology, pp. 209–226. Springer (2014)
11. Deladisma, A., Cohen, M., Stevens, A., Wagner, P., Lok, B., Bernard, T., Oxendine, C., Schumacher, L., Johnsen, K., Dickerson, R., et al.: Do medical students respond empathetically to a virtual patient? The American Journal of Surgery **193**(6), 756–760 (2007)
12. Draper, J.V., Kaber, D.B., Usher, J.M.: Telepresence. Human factors **40**(3), 354–375 (1998)
13. Finkelstein, S., Yarzebinski, E., Vaughn, C., Ogan, A., Cassell, J.: The effects of culturally congruent educational technologies on student achievement. In: International Conference on Artificial Intelligence in Education, pp. 493–502. Springer (2013)
14. Gerhard, M., Moore, D.J., Hobbs, D.J.: Continuous presence in collaborative virtual environments: Towards a hybrid avatar-agent model for user representation. In: International Workshop on Intelligent Virtual Agents, pp. 137–155. Springer (2001)
15. Gratch, J., Wang, N., Gerten, J., Fast, E., Duffy, R.: Creating rapport with virtual agents. In: International Workshop on Intelligent Virtual Agents, pp. 125–138. Springer (2007)
16. Heeter, C.: Being there: The subjective experience of presence. Presence: Teleoperators & Virtual Environments **1**(2), 262–271 (1992)
17. Hendrix, C., Barfield, W.: Presence within virtual environments as a function of visual display parameters. Presence: Teleoperators & Virtual Environments **5**(3), 274–289 (1996)
18. Hoffmann, L., Krämer, N.C., Lam-Chi, A., Kopp, S.: Media equation revisited: do users show polite reactions towards an embodied agent? In: International Workshop on Intelligent Virtual Agents, pp. 159–165. Springer (2009)
19. Juan, M.C., Pérez, D.: Comparison of the levels of presence and anxiety in an acrophobic environment viewed via hmd or cave. Presence: Teleoperators and Virtual Environments **18**(3), 232–248 (2009)
20. Kenny, P., Parsons, T.D., Gratch, J., Rizzo, A.A.: Evaluation of justina: a virtual patient with ptsd. In: International Workshop on Intelligent Virtual Agents, pp. 394–408. Springer (2008)
21. Kopp, S., Krenn, B., Marsella, S., Marshall, A.N., Pelachaud, C., Pirker, H., Thórisson, K.R., Vilhjálmsson, H.: Towards a common framework for multimodal generation: The behavior markup language. In: International workshop on intelligent virtual agents, pp. 205–217. Springer (2006)
22. Krämer, N.C.: Social effects of virtual assistants. a review of empirical results with regard to communication. In: Proceedings of the international conference on Intelligent Virtual Agents (IVA), pp. 507–508. Springer-Verlag, Berlin, Heidelberg (2008)
23. Krijn, M., Emmelkamp, P.M., Biemond, R., de Ligny, C.d.W., Schuemie, M.J., van der Mast, C.A.: Treatment of acrophobia in virtual reality: The role of immersion and presence. Behaviour research and therapy **42**(2), 229–239 (2004)
24. Lessiter, J., Freeman, J., Keogh, E., Davidoff, J.: A cross-media presence questionnaire: The itc-sense of presence inventory. Presence: Teleoperators and virtual environments **10**(3), 282–297 (2001)

25. Limniou, M., Roberts, D., Papadopoulos, N.: Full immersive virtual environment cave tm in chemistry education. Computers & Education **51**(2), 584–593 (2008)
26. Lok, B., Ferdig, R.E., Raij, A., Johnsen, K., Dickerson, R., Coutts, J., Stevens, A., Lind, D.S.: Applying virtual reality in medical communication education: current findings and potential teaching and learning benefits of immersive virtual patients. Virtual Reality **10**(3-4), 185–195 (2006)
27. Mestre, D.R.: On the usefulness of the concept of presence in virtual reality applications. In: IS&T/SPIE Electronic Imaging, pp. 93,920J–93,920J (2015)
28. Mikropoulos, T.A., Natsis, A.: Educational virtual environments: A ten-year review of empirical research (1999–2009). Computers & Education **56**(3), 769–780 (2011)
29. Monden, K., Gentry, L., Cox, T.: Delivering bad news to patients. Proceedings (Baylor University. Medical Center) **29**(1) (2016)
30. Parsons, T.D., Rizzo, A.A.: Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. Journal of behavior therapy and experimental psychiatry **39**(3), 250–261 (2008)
31. Pelachaud, C.: Studies on gesture expressivity for a virtual agent. Speech Communication **51**(7), 630–639 (2009)
32. Porhet, C., Ochs, M., Saubesty, J., Montcheuil, G., Bertrand, R.: Mining a multimodal corpus of doctor's training for virtual patient's feedbacks. In: Proceedings of 19th ACM International Conference on Multimodal Interaction (ICMI), Glasgow, UK (2017)
33. Quest, T.E., Ander, D.S., Ratcliff, J.J.: The validity and reliability of the affective competency score to evaluate death disclosure using standardized patients. Journal of palliative medicine **9**(2), 361–370 (2006)
34. Raij, A., Kotranza, A., Lind, D.S., Lok, B.: Virtual experiences for social perspective-taking. In: Virtual Reality Conference, 2009. VR 2009. IEEE, pp. 99–102. IEEE (2009)
35. Rosenbaum, M., Ferguson, K., Lobas, J.: Teaching medical students and residents skills for delivering bad news: A review of strategies. Acad Med **79** (2004)
36. Santos, B.S., Dias, P., Pimentel, A., Baggerman, J.W., Ferreira, C., Silva, S., Madeira, J.: Head-mounted display versus desktop for 3d navigation in virtual reality: a user study. Multimedia Tools and Applications **41**(1), 161 (2009)
37. Schubert, T., Friedmann, F., Regenbrecht, H.: The experience of presence: Factor analytic insights. Presence: Teleoperators and virtual environments **10**(3), 266–281 (2001)
38. Schubert, T.W.: The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. Zeitschrift für Medienpsychologie **15**(2), 69–71 (2003)
39. Shiratuddin, M.F., Thabet, W., Bowman, D.: Evaluating the effectiveness of virtual environment displays for reviewing construction 3d models. CONVR 2004 pp. 87–98 (2004)
40. Slater, M., Linakis, V., Usoh, M., Kooper, R., Street, G.: Immersion, presence, and performance in virtual environments: An experiment with tri-dimensional chess. In: ACM virtual reality software and technology (VRST), vol. 163, p. 72. ACM Press New York, NY (1996)
41. Slater, M., Sadagic, A., Usoh, M., Schroeder, R.: Small-group behavior in a virtual and real environment: A comparative study. Small-Group Behavior **9**(1) (2006)
42. Slater, M., Wilbur, S.: A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. Presence: Teleoperators and virtual environments **6**(6), 603–616 (1997)
43. Usoh, M., Catena, E., Arman, S., Slater, M.: Using presence questionnaires in reality. Presence: Teleoperators and Virtual Environments **9**(5), 497–503 (2000)
44. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: A presence questionnaire. Presence: Teleoperators and virtual environments **7**(3), 225–240 (1998)