Virtual Interactive Tablet to Support Vocational Training in Immersive Environment

Pierre Gac\textsuperscript{1,2}, Paul Richard\textsuperscript{1}, Yann Papouin\textsuperscript{2}, Sébastien George\textsuperscript{3} and Émmanuelle Richard\textsuperscript{1}

\textsuperscript{1}Université d’Angers, LARIS-EA 7315, 49000, Angers, France
\textsuperscript{2}DEC Industrie, 3 rue du Champ du Verger, 72200, ALLONNES, France
\textsuperscript{3}Le Mans Université, LIUM-EA 4023, 72085, Le Mans, Cedex 9, France

Keywords: Virtual Reality, Diegetic Interface, Vocational Training, Learning, Virtual Agent, Virtual Helper.

Abstract: This paper presents a tool designed to assist a virtual reality learner in a vocational training context. The benefits of VR are spreading outside the video-games and research field, leading the vocational education institutions to consider using this technology for training purposes. By simulating emblematic professional situations, teachers can train students in good conditions regarding safety, logistics and financial resources. In the French vocational training system, VR is just at its beginnings and the lack of experiments with this training context highlighted the need for new tools allowing teachers to use VR with their students. The work presented in this paper is part of a global project aiming to create, design and assess new VR tools and methodologies for this specific context. In order to guide, inform and assist the user, we are presenting a generic tool that can ease VR sessions by proposing embedded tools within easy reach of the immersed user, such as a camera, an inventory, a printer or stock management. This tool is a virtual tablet that can be grabbed by the immersed user allowing her/him to interact efficiently with the virtual environment.

1 INTRODUCTION

The emergence of commercial headsets such as HTC Vive\textsuperscript{TM}, Oculus Rift\textsuperscript{TM} or smartphone-based solutions, highlights the possibilities of using VR technology for a large public. Nowadays, nearly everyone has heard about VR. The original target of commercial VR devices is video games, that is why when people think about VR, they are instantly picturing someone playing games with a headset and wireless controllers. Since the arrival of the new line of commercial headsets, the global perception of VR has been evolving leading to new opportunities of VR usage in other fields such as education. In our research projects, we are using VR devices for vocational training. Using modern VR on this field is something relatively new, both teachers and developers suffer from the lack of perspective about VR applied for vocational training, which implies a new development approach in order to create an effective learning environment.

Vocational training is a very specific field in education with different ways and practical applications across the world. That is why some studies about VR and virtual learning cannot be applied to French vocational education. Since a few years, the French government is pushing forward digital transition in education, and VR is now considered as a new training support. In the case of French vocational training, we are facing different problems that VR can pretend to fix. In some sectors, there is a huge need of expensive equipment and not all high schools can claim that equipment. Some French academies have more financial and logistical support than others, which causes great disparities in training offers. The use of VR for these specific situations will make it possible to propose virtual but realistic environments to provide access to these important and expensive equipment and thus, allow students to practice in an immersive way. The other related problem is that in vocational training, students can make mistakes that can have negative consequences on equipment or endanger the safety of people. However, these errors can be useful for the learning process. VR can be a solution, allowing students to try without consequences in the real world. Furthermore, the teacher can use these mistakes to give formative feedback to students during debriefing sessions. The other motivation of using
VR in vocational training is the safety and autonomy of students.

Vocational training covers various fields such as electrical maintenance or sales. While developing immersive training software, the designer is tempted to setup realistic interactions in the same way as video-games or marketing purposes. VR can show realistic environments and interactions, but fully imitating the reality in VR will question the relevance of this technology for training purposes, also it is sometimes interesting to add unrealistic behavior to simplify tasks or shorten unnecessary tasks. The aim of this specific simplification is to attract the student’s attention over a precise goal. Choosing relevant tasks and interactions drive us to design a versatile diegetic interface that allows to use some tools intended to simplify the virtual tasks. During the development, we have also taken into account the diversity of vocational training situations. In this paper we will expose our tool in the specific context of the French vocational training system. Our research projects are focused on VR design and pedagogical relevance of VR applied to vocational training in order to allow teachers to use them easily in their learning programs.

The contribution of this paper is to propose a virtual interactive tablet that is designed to help and assist students to accomplish the virtual tasks. To aim for this, we developed a virtual tablet that can be grabbed by the user in the virtual world and use it as a toolbox or as a helper. We will first make a review of the literature related to Pedagogical Virtual Agents (PVA) and 3D user interfaces (UI). This review of the literature is meant to explain the link between a PVA and UI. Then, we will add more details about our proposal and the context of this work. We have conducted some preliminary tests and the results will be exposed in a specific section. Finally, we will conclude with the ongoing experiments that we are planning to test our tool as well as other possible iterations of the virtual tablet.

2 RELATED WORK

2.1 Pedagogical Virtual Agents

In VR, we sometimes need to guide the user through the environment and goals. To do so, several techniques were created such as PVA. Those embodied agents are virtual avatars which can have multiple functions or representation. The first iterations of PVA highlighted the need of a helper that can guide and also show manipulation examples by its interaction capabilities to the user. The virtual agent will react to user’s actions and provide specific assistance. Such interaction is a good vector for student motivation thanks to customized advices. An agent persona can be described using the four-factor model as detailed in Ryu and Baylor (2005) (facilitating, engaging, credible and human-like). In Jin (2010), students found the VE with PVA more entertaining and educational than the one without it. Besides reacting to interactions (object manipulation, user’s action or environment events), virtual agents can handle conversational capabilities as described in the MAX framework (Kopp et al., 2003; Lopez et al., 2014). Their solutions can react to user’s questions or affirmations. This kind of conversational assistant is now present in today’s life with assistants like Siri, Google Home or Alexa. It is now natural for someone to give a speech to a computer.

With the emergence of machine learning and AI, PVA can now mimic a more realistic behavior and can follow the student’s progression. To do so, agents have to react to specific situations with intelligence and adequate behavior as described by Mittal et al. (2013) to have a believable behavior. To build an efficient and credible behavior of the agent, the agent system needs to have access to all world/situation knowledge (omni-data) but it has to limit its access to this omni-data otherwise it will result in unrealistic behavior (Mittal et al., 2013). It is also possible to make the agent aware of the real-world situation, thanks to new sensory devices, given that, a real world-aware agent can adapt its speech, gesture, or facial expression according to the user in real-time (Wuttke et al., 2016). In addition, decision making is a key element to the realistic behavior of the PVA as stated Lopez et al. (2014). VR is a conducive mean to collaboration, this could be achieved with multiple humans or one human and several agents. In their HUMANS suite, Lourdeaux et al. (2017) have implemented both individual and collaborative capabilities to their agents.

2.2 3D User Interface

In classic 2D games, the user interface (UI) area is displayed on the camera view (overlay mode), while in VR, we can setup a UI that is present in the 3D world. Our work is focusing on those 3D menus that use widgets and classic UI. Voice and gestures UI are not compatible with our work context (vocational training). For voice command UI, we faced three issues that drove us to drop this approach. First, the class organization around the VR system can generate noise, making the system unable to correctly interpret oral commands. Secondly, we want to take into account the speaking manners of some students (e.g.
strong accents, or language disorders). The third point is about technical limitations such as programming and logistics (e.g. inputs or Internet connection). For the gesture UI, we decided not to consider this manner of interaction because it will involve more development time than the "classic" UI as well as for the user having to memorize the gesture patterns. Many types of 3D UI selection metaphors have been developed for VR applications. For instance, Bowman et al. (2001) proposed a UI system that is controlled by the user’s real hands, using gloves. Before commercial devices like HTC Vive™, 3D UI interaction were made with several unique devices. They also used a tablet-based interaction that allows the user to select a virtual floating menu item by using a real tablet. Those interactions are not efficient with today’s tools, for example, with the HTC Vive™, the user has an efficient tracking of its hands in the virtual world allowing specific interactions. This allows the user to easily interact with 3D UI with some feedback such as the vibrations of controllers, sounds or visuals.

In vocational VR training, we are looking for pseudo realistic situations where the learner can be projected in order to find a link between the virtual experience and the real one. Sometimes in VR we need to display textual information, the problem for the designer is to choose the appropriate metaphor for displaying text. The diegesis theory is initially a literature concept, that has been applied to videogames and thus, to VR by Gallioway (2006). Finally, the VR designer can use four kinds of UI: Diegetic, non-diegetic, spatial or meta as described in Fagerholt and Lorentzon (2009). The diegetic interfaces are strongly integrated in the virtual world (Figure 1). They proposed the classification regarding 2 conditions: “Is the representation visualized in the 3D game space?” and “Is the representation existing in the fictional game world?”. The answers are crossed into an array. In a sci-fi universe, a diegetic UI can be an holographic representation of the menu. In FPS, it can be the way ammo are displayed (Peacocke et al., 2016) or health displays. The majority of studies about the effectiveness of diegetic UI are empirical, and have been tested by some videogame studios.

A positive aspect of diegetic interfaces, is that they tend to improve feelings of immersion (Raffaele et al., 2017). In their work, Raffaele et al. (2017), investigated the relations between immersion and virtual interfaces through VR. Their results indicate that the user will receive feedback from the UI without being conscious about it and not noticing how feedback is received. For 3D UI, not noticing the feedback source is pointed out as the key to keeping the immersion level high. Consequently, in our training context, it can benefit the learner. Therefore, the design of such UI is meaningful in order to not confuse the final user. They conclude by saying that Spatial and Diegetic UI are the most immersive types of UI. To achieve a correct immersion level, unless considering the quality of the environment, VR designers have to favor Spatial or Diegetic UI to keep immersion at the top most levels. Feelings of immersion are sometimes disrupted by the controllers complexity (Fragoso, 2014). In the case of external UI, the feelings of immersion do not vanish if the UI maintains the aesthetic perception (Fragoso, 2014).

Llanos and Jørgensen (2011) suggests that a minimalist interface is easier to manipulate and so, provides better information as it is easier to access and process. Participants said that minimalism is an "elegant way of representing system information". To consolidate the importance of minimalist UI, Gerber and Bechmann (2004) noticed that the average selection time grows with the number of items displayed. It should point out that today’s standards of UI in smartphones or computers are moving forward to the minimalist approach as it is considered more "user-friendly". Llanos and Jørgensen (2011) add in their conclusion that users prefer to receive only relevant information instead of a huge quantity of data. Additional information that is not necessary will be tiresome and thus, will somehow break the immersion and force the learner to focus more on the UI. The precision of the data displayed is a sensitive point for VR designers.

3 CONTEXT OF THIS WORK

This work is part of a large-scale project that consists of creating VR applications for vocational training. The French Education Academy of Nantes is the pilot academy in VR experiment and deployment. We are working with teachers and inspectors on the next iteration of digital training using VR. The skill-based
approach implies new design processes regarding scenario writing, software development and evaluation. There is a significant need of "useful" emblematic training situations that can be used by the teacher and professionals.

This project is divided in three main parts. The first part is the scenario design methodology and "in class" use of VR as well as sharing our results and findings to the teachers in a way to improve the use of VR in an Educational context. Existing methodologies presented in the VR literature cannot directly be applied to our design approach. These methodologies are arduous to understand for teachers and professionals. The second part relates to the VR development, this is about creating a SDK that allows the user to create a VR training scenario from scratch with Unity3D without the need of consequent programming skills. The targeted user can be a confirmed developer or a teacher. The third part deals with students’ assessment by the VE and by the teacher. Vocational training in France has its own specificities that we need to take into account to build efficient training situations. Didactic knowledge is the key to achieve a good implementation of VR training scenario.

4 VIRTUAL AGENTS FOR VOCATIONAL TRAINING

In vocational training, users need to be guided while practicing. Students require additional information and help to achieve tasks they are not mastering. While progressing and practicing, learners will gain autonomy and will need less external help to complete tasks. Using VR will not decrease this need of help, instead, students will need more help to master the tools and scenarios. That is why, when composing a VR training application, designers should include a virtual assistant system, embodied or not.

In our project, we first considered using embodied PVA as described in the literature for the benefits it provides (Kosinowski, 2009). Nevertheless, in vocational training some situations are not compatible with the presence of PVA. For electrical maintenance or production line management, it is uncommon but sometimes dangerous to have someone near the repair task. The area of work can be too small, or standard tasks do not allow more than one person in the work area (e.g., scaffold inspection, electrical clearance). In that case, the PVA should be outside the activity space, which implies the user to navigate from and to the work area to interact with the PVA. The type of agent for vocational training purposes could be an instructor as described by Sklar (2003). The problem with instructors is that over-guiding the user can "get in the way of the user’s learning".

The other concern about PVA for us, is that virtual agents will require huge development time and resources to achieve realistic behavior (e.g., autonomy, interactions, intelligence) as described in the literature. Moreover, developing a PVA requires a significant amount of time for refinement (Barthes et al., 2018). The user need to trust the agent and to achieve that reliability is not enough, attractiveness is just as important as reliability (Yuksel et al., 2017). Therefore, implementing trusting agent will cause more design work. A second-rate development like poor animations, poor facial expression or low-quality sounds can lead to an awful user experience and poor attractiveness. We are using virtual agents only for simple dialogs like "Take this document" or "You need to do those tasks". These are presented as a dialog popup near the agent’s avatar and also played with sound speech. This is a requirement for the students who have difficulties at reading and hearing. The consideration of disabilities such as hearing impairment or difficulty of reading is not always done in VR software. In addition of this easy use of PVA, we still need a tool that improves the users’ autonomy and allows them to get help or advice when practicing in VE. The idea of an interface that encapsulates all of these additional functionalities came through.

5 THE VIRTUAL TABLET

![Figure 2: The home menu of the tablet.](image)

The constraints encountered while creating the VR training application drives us to consider new ways of guiding and assisting the user. This idea came from the observation of some virtual tasks. Some tasks like printing a document have low pedagogical interest and cost a huge amount of time. Indeed, the user need to go to the printer, print the document and then return to the right location to continue the task. A floating 3D UI would appear above the printer. The learner had to select the document to print and then
click on a "print" button. After doing that, they had to go back to the activity area. This time-consuming process was inefficient because if the student had forgotten something or printed the wrong document, s/he had to navigate back to the printer area, while wasting time traveling. By the end of the scenario, the student will have been disinterested of navigating through the VE. In vocational training, these kinds of errors are very common, because students tend to forget or not read the activity instructions.

Another problem encountered was the document reviewing in VR. The naive approach was to copy the reality into the virtual. Teachers initially thought that if they copy reality inside the virtual world, the student will be able to perform tasks the same way as in the real world. They did not consider the interaction process/tools, development complexity and the user’s experience. They asked to have directions, plans, or documents printed on virtual A4 paper displayed inside the virtual world. They were confident over the results, but after testing this, they noticed that the documents were unreadable and required a lot of attention and focus to be fully understandable. This led us to propose another way to visualize documents, by removing the virtual A4 paper for the benefit of a minimalist 3D UI with only necessary information displayed with better font and contrast.

VR in vocational training is not meant to mimic the reality at 100%, because in that case, why would VR be useful? The good exploitation of VR is a requirement to build adequate training situations. To do so, we need the following elements:

- Access to directions at anytime
- Simple interactions (to understand and to use)
- A solution that can be used in several training situations (generic-programming)
- Simplify real interactions by removing irrelevant tasks/steps
- Give/Gather information from/to the user
- Display warning or directions at a specific time when needed
- Guide the user through the scenario

With all of those elements, and considering the case of PVA, we came up with the idea of the tablet metaphor solution. The concept is to build a generic tool that encapsulates all of those elements and more. The interesting point of the tablet metaphor is the equivalence with real emblematic situations. In a lot of working fields, workers are using tablets, smartphones or PDA in their tasks (e.g. sales, electrical maintenance, logistics, building workers). That is why using a virtual tablet as a model for the interface is relevant because of the similarities of the virtual device to the real world use of devices even when our virtual content does not exactly match the real use of those tools.

This device is held by the user using one of her/his hands and is considered by our SDK as a passive device. It means that the virtual hand containing the tablet can not interact with the world by raycasting scene objects. The student can switch the hand holding the tablet by simply clicking on the tablet with the other hand. The aim of locking the tablet on one hand is to allow the student to access it anytime and avoid losing the tool somewhere in the VE. The purpose of VR training situation is not to search for a tool or to wander in the VR. While testing our VR solutions, we noticed that wandering can lead the student to not fully understand what teachers expect from her/him or can feel that they are lost in the scenario. With the tablet, we can mostly eliminate all potential wandering situations.

For the document printing process, we achieved to eliminate the unnecessary navigation part, the printing station interactions and reduced possible errors that can be made by the user because s/he can print documents near the activity area. To achieve all that, we put this process inside the tablet (see Figure4). An application called "Printer" was created, in which the user can select a document from a category and print it. As Fragoso (2014) established, clarity and functionality are more important than integration in the game world, so we decided to make a simple interface for this tablet application. The user can select a document by clicking on an image and then click on "print". The printed document will appear at the top of the tablet and the student just have to select it like she/he would do for any virtual object. This part is not realistic, but it does not negatively impact the training process, as the methodology behind document printing is the same regardless of the interaction. All of this process is achieved without any unnecessary navigation in the VR world. We use the same approach for other interactions such as the document reading process, the inventory or directions.

We implemented a "home" screen (Figure2) where the user can start an application by simply clicking on the corresponding icon. We also chose to include applications that can be used anytime in any training situation, such as the camera that allows the user to take pictures of the game world. The images taken can be retrieved by the user after the virtual sessions. This "simple" tablet application has a huge training interest as teachers can ask students to capture things that they judge "interesting". The captured images can be used during a debriefing step or
in other courses like written expression. In the case of a shop safety inspection, teachers ask the students to explain why they took this photo by writing down their analysis process.

One advantage of VR is data gathering (traces). We are currently using the Bowman et al. (2001) classification to sort out the interactions made by the user: manipulation, selection, navigation and system control. The tablet is too advanced to be considered as a simple manipulation or selection. The system control is the category that is mostly compatible with the tablet. Bowman defines the system control as "a task in which a command is applied to change either the state of the system or the mode of interaction". The tablet behaves in a similar way as described by Bowman, but some applications cannot be included in the system control. For example, the camera application does not change the mode of interaction or the state of the system. It is just a selection in this case, on the contrary to the inventory application that will change the mode of interaction when selecting an object. Thus, the tablet includes all types of interactions, and a deeper investigation is needed to find a correct categorization of the device. A potential category may be "Functionality selection" as the tablet allows the user to select several features.

6 PRELIMINARY EXPERIMENTS

To prepare our system for the planned experiments, we have conducted two preliminary tests using the HTC VIVE™. The chosen activity was an order taken from the application (completion time, amount of correct/miss clicks, name of clicked objects, current application) as well as with a questionnaire. We used a French translated version of the System Usability

6.1 Preliminary First Experiment

The first experiment, involved high-school students (6 participants) in sales training. During those sessions, teachers and us gathered subjective feelings of students regarding the virtual situation and the tablet. In addition to our research purposes, teachers wanted to assess if VR can be an interesting tool for several student’s profiles. One noticeable observation we made is that users quickly master the tablet interface and the purpose of each application. According to students, it is thanks to the "close design" of the tablet to real devices like Android systems. Navigation on the tablet seems to be "natural" thanks to explicit images, concise labels and good UI contrasts. We also asked the question "Is the tablet useful in this scenario?", to what participants answered "the tablet is useful because we can quickly access stuff, like the printer or the stock screen, without loosing time doing like in the real world". "With the tablet, we are still focused to what we are doing and not tempted to wander and play like in a video game". One student also added that if the scenario had transposed the real world into the virtual (without the tablet), then "the situation would be annoying and so does VR for training". Users agreed that the tablet is an efficient assistant and grant them autonomy in the virtual scenario. Besides those positive feedbacks, we noticed that users sometimes strangely handle the tablet, by keeping it in front of their view instead of lowering the device to clear the view. They did not notice this until we say "You do not need to tablet now, you can lower your arm if you want". We have also observed issues with some UI design regarding the size of some buttons and labels. Two users reported that handling the tablet on one hand can be exhausting over the time, that is why we may need to consider new interaction methods, different than the actual raycasting, for long VR activity (more than 20min). We observed that the raycasting is sometimes hard to achieve especially on the top panel, because this panel is too small. According to teachers, the virtual tablet is interesting because "It is similar to some work situations where workers use technology to ease their jobs". This interest is also present in the student point of view.

6.2 Preliminary Second Experiment

The second experiment was conducted with adults (7 participants). The main purpose was to gather data from the application (completion time, amount of correct/miss clicks, name of clicked objects, current application) as well as with a questionnaire. We used a French translated version of the System Usability
Scale (SUS) (Brooke and Weerdmeester, 1996) and replaced the word “system” by “tablet”. The mean score is 79 (σ=13). From traces gathered, we noticed that users perform a small amount of miss clicks on the tablet. A miss click occurs when the user is clicking on a non interactive component like label or image instead of a button. For an average of 111 (σ=32) clicks on the tablet per session, users performed 2.8% miss clicks (σ=3%). For this second test, we have changed some fonts on the tablet, because we noticed in the first test sessions that users with difficulty of reading had trouble to read instructions on the tablet because of the font we initially choose (Oswald Regular). We tried a dyslexic friendly font (OpenDyslexic), resulting in poor results in the CAVE and headsets, mainly because of the resolution. Afterwards, we chose the Verdana font for the second test and users do not pointed out any text issues. Data showed that the tablet is not a source of difficulty, and user’s profile does not impact tablet usage. Furthermore, objective data retrieved from the tablet cannot be used alone to define user’s profile and categorize it.

7 DISCUSSION

The usability of the tool relies on the UI clarity and design, we chose a minimalist approach with only relevant information, to not overload the student with much many information. The two first iterations faced major design issues as well as generic programming problems. We are still facing a major issue regarding the degree of realism to include in course of the pedagogical virtual scenario. Indeed, we noticed that a minimalist design is more welcomed by students from initial formation than by teachers or older students. Those young students appear to be more comfortable in the VE than teachers. Teachers want to display as much information as possible without considering the readability problems in VR. We tried this approach and the result is that students were lost because of the overwhelming quantity of information. With the minimalist UI, we need to train and convince the teachers that we can not display a large amount of information because it will have a negative impact over students’ performance. The initial wish of teachers is to take all the information from the real world and transpose it into the virtual. As explained earlier, this is a bad design choice and will cause bad results.

The tablet is held by one of the user’s hand, so the hand containing the device can not be used to perform other interactions. The user will be able to interact with only one hand. In some training situations, it is an issue. For example, a voltage tester will require two hands to operate. So, the user has to place down the tablet somewhere, perform the action and then grab the tablet again. We are investigating another metaphor, instead of having the tablet on one hand, the device will be placed on the arm of the user, allowing her/him to interact with their two hands. This new approach is also facing some issues. The first issue is the size of the screen, and for it to be credible we need to reduce its size. Downsizing can have a negative impact over the readability of the device and thus, its usability. Another issue is the interaction metaphor used to interact with the arm-based tablet. We are actually using a raycast from the hand that is not containing the tablet to do the selection process. With an armed-based solution, performing a raycast is practically complicated and requires the user to contort itself to select an item on the arm-based tablet. For this case, other metaphors are considered, like a hybrid grab: the interaction is active when the interactive object is within a certain range around the virtual hand.
8 CONCLUSION AND FUTURE WORK

This paper presents the design of a training-oriented tool that allows the user to manage several interactions with only one tool and without involving a complex interaction process to achieve the desired actions. To assess our design choices, we are planning experiments with teachers and students. We are interested in student's performances with or without the tablet as well as their feelings in the training environment. Some non-realistic interactions from the tablet will have to be implemented in a realistic way in the VE (like the printing station described earlier). In addition to the immersive data gathered we will ask teachers to assess student performances with the data from the tablet and with data from a situation without the tablet. We also want to experiment other metaphors such as the arm-based solution previously described.

REFERENCES


