Structuring Interactions in a Hybrid Virtual Environment
*Infrastructure & Usability**

Pablo Almajano[1,2], Enric Mayas[2], Inmaculada Rodriguez[2], Maite Lopez-Sanchez[2] and Anna Puig[2]
[1]*Institut d'Investigació en Intel·ligència Artificial, CSIC, Bellaterra, Spain*
[2]*Departament de Matemàtica Aplicada i Anàlisi, Universitat de Barcelona, Barcelona, Spain*

Keywords: Virtual Environments, Human-agent Interactions, Usability.

Abstract: Humans in the Digital Age are continuously exploring different forms of socializing on-line. In Social 3D Virtual Worlds (VW) people freely socialize by participating in open-ended activities in 3D simulated spaces. Moreover, VWs can also be used to engage humans in e-applications, the so called Serious VWs. Implicitly, these serious applications have specific goals that require *structured environments* where participants play specific roles and perform activities by following well-defined protocols and norms. In this paper we advocate for the use of *Virtual Institutions* (VI) to provide explicit structure to current Social 3D VWs. We refer to the resulting system as *hybrid* (participants can be both human and software agents) and *structured* Virtual Environments. Specifically, we present *v-mWater* (a water market, an e-government application deployed as a VI), the infrastructure that supports participants' interactions, and the evaluation of its usability.

1 INTRODUCTION

Social 3D Virtual Worlds (VW) are a relatively new form of socializing on-line. They are persistent Virtual Environments (VE) where people experience others as being there with them, freely socializing in activities and events (Book, 2004). However, VWs can also be used to engage humans in serious applications with reality-based thematic settings (e.g., e-government, e-learning and e-commerce), the so-called Serious VWs.

Social VWs are conceived as *unstructured environments* that lack of defined and controlled interactions, whereas Serious VWs can be seen as inherently *structured environments* where people play specific roles, and some activities follow well-defined protocols and norms that fulfil specific goals. Although users in such structured (regulated) environments may feel over-controlled, with most of their interactions constrained, this can turn around and feel more guided and safe whenever regulations direct and coordinate their complex activities. Current VWs platforms (e.g Second Life), mainly focused on providing participants with open-ended social experiences, do not explicitly consider the definition of structured interactions, neither contemplate their control at run-time.

Therefore, we advocate the use of *Virtual Institutions* (VI), which combine *Electronic Institutions* (EI) and VWs, to design *hybrid and structured Virtual Environments*. EIs provide an infrastructure to regulate participants' interactions. Specifically, an EI is an organisation centred Multi-Agent System (MAS) that structures agent interactions by establishing the sequence of actions agents are permitted/expected to perform (Esteva et al., 2004). VWs offer an intuitive interface to allow humans to be aware of MAS state as well as to participate in a seamless way. By *hybrid* we mean that participants can be both human and bots (i.e., software agents). They both perform complex interactions to achieve real-life goals (e.g., tax payment, attending a course, trading).

In this paper we present an example of a hybrid regulated scenario in an e-government application (*v-mWater*, a *v*irtual *m*arket based on trading *Water*), the infrastructure that supports participants' interactions in this scenario, and the evaluation of its usability. The *Virtual Institutions eXEcution Environment* (VIXEE) (Trescak et al., 2011) infrastructure enables the execution of a VI. As far as we know, there are no previous evaluations about the usability of an application deployed with a similar infrastructure (i.e. a strongly regulated and hybrid virtual environment). We are specially interested in analysing how users perceive their interaction with bots.

*Work funded by EVE (TIN2009-14702-C02-01/02), at (CONSOLIDER CSD2007-0022) and TIN2011-24220 Spanish research projects, EU-FEDER funds.

Almajano P., Mayas E., Rodriguez I., Lopez-Sanchez M. and Puig A..
Structuring Interactions in a Hybrid Virtual Environment - Infrastructure & Usability.
DOI: 10.5220/0004215802880297
In *Proceedings of the International Conference on Computer Graphics Theory and Applications and International Conference on Information Visualization Theory and Applications* (GRAPP-2013), pages 288-297
ISBN: 978-989-8565-46-4

2 RELATED WORK

In this section we review prior research works on structured (regulated) interactions in Multi-User/Agent environments, infrastructures that extend basic functionalities of VW platforms, and usability evaluation of virtual environments (VE).

Regulation has been subject of study both in Multi-Agent Systems (MAS) and Human Computer Interaction (HCI) fields. In the MAS field, several studies focused on agents societies and proposed methodologies and infrastructures to regulate and coordinate agents interactions (Dignum et al., 2002) (Esteva et al., 2004). Specifically, Cranefield et al. adapted a tool, originally developed for structuring social interactions between software agents, to model and track rules of social expectations in Second Life (SL) VW such as, for example, “no one should ever fly” (Cranefield and Li, 2009). They used temporal logic to implement the regulative system. In our case, we use Electronic Institutions (EI), a well known Organization Centered MAS, to regulate participants' interactions in hybrid 3D VEs.

Several HCI researches focused on regulation mechanisms for groupware applications, i.e CSCW (Computer Supported Collaborative Work). In general, these mechanisms define roles, activities and interaction methods for collaborative applications. One research work used social rules (and the conditions to execute them) to control the interactions among the members of a workgroup (Mezura-Godoy and Talbot, 2001). Another work proposed regulation mechanisms to address social aspects of collaborative work such as the location where the activity take place, collaborative activities by means of scenarios, and the participants themselves (Ferraris and Martel, 2000). At a conceptual level our regulation model based on EIs (i.e activities, protocols, roles) shares similarities with those applied for groupware applications.

Related to regulation of activities in VEs, Paredes et al. proposed the Social Theatres model. This model regulates social interactions in a VE based on the concept of *theatre* (i.e. a space where actors play roles and follow a well-defined interaction workflow regulated by a set of rules). In posterior works, they conducted a survey to evaluate user preferences about VE interfaces. This allowed to design a 3D interface based on the Social Theatres model and users' preferences (Guerra et al., 2008). Recently they have proposed a multi-layer software architecture implementing the Social Theatres model (Paredes and Martins, 2010). Although it has been designed to be adaptable, this architecture presents some limitations on the dynamic adaptation of rules. On the contrary, as long as our system uses an EI as regulation infrastructure, it inherits self-adaptation properties of EIs (Campos Miralles et al., 2011). Another main difference between our system and the Social Theatres model is that the latter is a web-based environment (relying in web services) and our system is independent of the technology that implements the VE.

There is a variety of works that have connected multi-agent systems to VW platforms. CIGA (van Oijen et al., 2011) is a general purpose middleware framework where an in-house developed game engine can be connected to a MAS. Another middleware was proposed as a standard to connect MAS systems to environments (Behrens et al., 2011). They proposed a so called Environment Interface Standard (EIS) which supports several MAS platforms (2APL, GOAL, JADEX and Jason) and different environments (e.g. GUI applications or videogame worlds). The infrastructure that we present in this paper, regulates participants' interactions at run-time and provides the virtual space with intelligent behaviors.

There are also recent research works that have focused on extending, using plug-ins, VW platforms with advanced graphics for serious applications. As an example, a framework for Open Simulator creates scientific visualizations of biomechanical and neuromuscular data which allows to explore and analyse interactively such data (Pronost et al., 2011).

Regarding usability evaluation of VEs, Bowman et al. analyzed a list of issues such as the physical environment, the user, the evaluator and the type of usability evaluation, and proposed a new classification space for evaluation approaches: sequential evaluation and testbed. Sequential evaluation, that includes heuristic, formative/exploratory and summative evaluations, is done in the context of a particular application and can have both qualitative and quantitative results. Testbed is done in a more generic evaluation context, and usually has quantitative results obtained through the creation of testbeds that involve all the important aspects of an interaction task (Bowman et al., 2002). There are a number of researches that have proposed different evaluation frameworks for collaborative VEs (Tsiatsos et al., 2010) (Tromp et al., 2003). The approach that we have followed in this research paper is the sequential approach, mainly formative because we observe users interacting in our hybrid environment but also summative because we take some measures of time and errors performing tasks.

3 EXAMPLE SCENARIO

Our example scenario is a *v*irtual *m*arket based on trading *Water* (*v-mWater*). It is a simplification of m-Water (Giret et al., 2011) implemented as a VI which models an electronic market of water rights in the agriculture domain (Almajano et al., 2012).

3.1 Water Market

In our market, participants negotiate water rights[1]. An *agreement* is the result of a negotiation where a seller settles with a buyer to reallocate (part of) the water from her/his water rights for a fixed period of time in exchange for a given amount of money.

We consider farmlands irrigating from controlled water sources within a hydrographic basin. Public authorities estimate water reserves and assign a given water quantity to each water right. Irrigators that estimate they will have water surplus can then sell their rights. Our market only allows to enter and participate in the negotiation irrigators holding rights, i.e. farmlands, in the hydrographic basin.

3.2 Specification of Interactions

We use an Electronic Institution (EI) to structure participants' interactions in the virtual environment (VE). An EI is defined by the following components: an *ontology*, which specifies domain concepts; a number of *roles* participants can adopt; several dialogic *activities*, which group the interactions of participants; well-defined *protocols* followed by such *activities*; and a *performative structure* that defines the legal movements of *roles* among (possibly parallel) *activities*. More specifically, a *performative structure* is specified as a graph where *nodes* represent both *activities* and *transitions* and are linked by directed *arcs* labelled with the roles that are allowed to follow them.

In the *ontology* of our water market scenario we have included concepts such as *water right*, *land* or *agreement*. Moreover, participants (both software agents and humans) can enact different roles. Thus, a *buyer* represents a purchaser of water rights, a *seller* is a dealer of water rights, a *market facilitator* is responsible for each market activity, a *basin authority* corresponds to the legal entity which validates the agreements, and an *institution manager* is in charge of controlling access to the market. To enter the institution, an agent must login by providing its name and the role it wants to play. Successfully logged-in agents are located at a default initial activity. From this activity, agents in *v-mWater* can join three different dialogical activities: in the *Registration* activity water rights are registered to be negotiated later on; in the *Wait&Info* activity, participants communicate each other to exchange impressions about the market and obtain information about both past and next negotiations; and finally, the negotiation of water rights takes place in the *Auction* activity. It follows a multi-unit Japanese auction protocol, a raising price protocol that takes, as starting price, seller's registered price. Then, buyers place bids as long as they are interested in acquiring water rights at current price.

[1] In an agricultural context, a water right refers to the right of an irrigator to use water from a public water source (e.g., a river or a pond). It is associated to a farmland and the volume of its irrigation water is specified in m^3.

Figure 1: Initial aerial view of v-mWater with three rooms (activities).

Participants and specification elements of an EI have their corresponding representation (visualization) in the 3D VE. As an example, participants are represented as avatars whereas activities are depicted as rooms with doors in order to control the access (see Fig. 1). Next section focuses on the infrastructure that supports such structured 3D VE.

4 INFRASTRUCTURE

We have used VIXEE, the *Virtual Institutions eXEcution Environment* (Trescak et al., 2011), as a robust infrastructure to connect an *Electronic Institution* (EI) to different *Virtual Worlds* (VW). It allows to validate those VW interactions which have institutional meaning (i.e contemplated in the EI specification), and update both VWs and EI states to maintain a causal dependence. It also contemplates the dynamic manipulation of VW content. This section describes the 3 layered VIXEE architecture depicted in Fig. 2.

4.1 Architecture

The Normative Control Layer. (NCL) on the left side of Fig. 2 is in charge of structuring interactions. It is composed by an *Electronic Institution Specifica-*

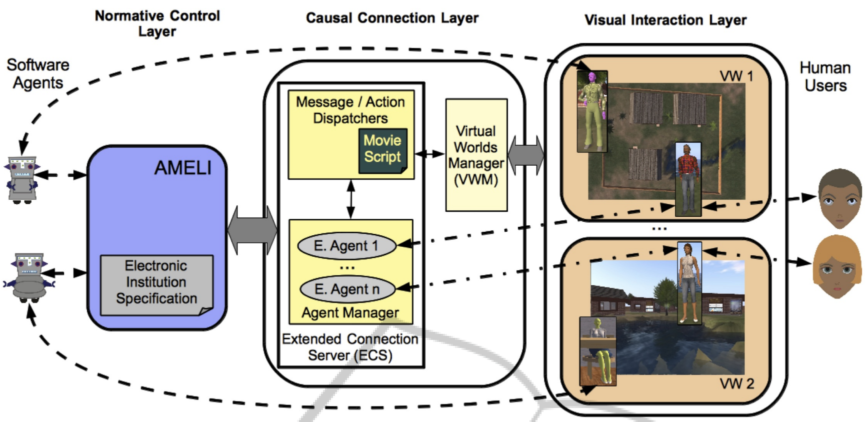

Figure 2: VIXEE Architecture. The Causal Connection Layer as middleware between the Normative Control Layer (populated by agents) and the Visual Interaction Layer (populated by 3D virtual characters).

tion[2] and AMELI (Esteva et al., 2004), a general purpose EI engine.

AMELI interprets such a specification in order to mediate and coordinate the participation of every agent within the MAS system.

Software (SW) agents (robot-alike icons on the left of Fig. 2) have a direct connection to AMELI. In turn, AMELI sends its messages to the middleware and receives messages from it describing VW actions.

The Visual Interaction Layer, represents several 3D VWs. Human users (human-face icons on the right of Fig. 2) participate in the system by controlling avatars (i.e. 3D virtual characters) which represent them in the virtual environment. Additionally, SW agents from the NCL can be visualised as bots in the VW (notice how dashed arrows in Fig. 2 link robot icons on the left with bot characters within this layer).

This layer may host VW platforms programmed in different languages and using different graphic technologies. The common and main feature of all VW platforms is the immersive experience provided to their participants.

VWs can intuitively represent interaction spaces (e.g. rooms) and show the progression of activities that participants are engaged in. For example, an auction activity can be represented as an room where the auctioneer has a desktop and dynamic panels show information about the ongoing auction. In order to explore the VW, users can walk around as done in real spaces, but they can also fly and even teleport to other places in the virtual space. Participant interactions can be conducted by using multi-modal communication (e.g. text chat, doing gestures or touching objects). The immersive experience can be still enhanced by incorporating sounds (e.g. acoustic signals when determining a winner in an auction).

The main components of this layer are VWs. We contemplate VW platforms based on a client-server architecture, composed by a VW client and a VW server. The former provides the interface to human participants. It is usually executed as a downloaded program in the local machine (e.g. Imprudence) or as a web interface. The latter communicates with the Causal Connection Layer (see in next section) by using a standard protocol (e.g. UDP, TCP, HTTP). In particular, the scenario described in § 5 employs *Open Simulator* (http://opensimulator.org), an open source multi-platform, multi-user 3D VW server.

The Causal Connection Layer, in Fig. 2 -or middleware- keeps a state-consistency relation: it connects human participants from multiple VWs to the NCL so to regulate their actions; and it supports the visualisation of SW agent participants as bots in the VWs. Next, we explain its main components.

The *Extended Connection Server* (ECS) mediates all the communication with AMELI. It supports the connection of multiple VWs to one EI. This way, users from different VWs can participate jointly in the same VI. Moreover, ECS is able to catch those AMELI messages that trigger the generation of the initial 3D environment (e.g. build rooms) and reset the world to a pre-defined state (e.g. clear information panels) [3]. ECS main elements are the *Agent Manager* and *Message/Action Dispatchers*. First, the *Agent Manager* creates an External Agent (E. Agent in

[2] In order to generate an *EI specification* we use ISLANDER, the EI specification editor that facilitates this task.

[3] ECS manipulates VW content by means of two components: the *Builder* and the *VW Grammar Manager*.

Fig. 2) for each connected (human-controlled) avatar. The E. Agent is connected to the EI with the aim of translating the interactions performed by the human in the VW[4]. Second, *Message/Action Dispatchers* mediate both AMELI messages and VW actions. They use the so called *Movie Script* mechanism to define the mapping between AMELI messages and VW actions and vice versa. On one hand, a message generated from AMELI provokes a VW action so that the visualisation in all connected VWs is updated. On the other hand, for each institutional action performed by a human avatar in the VW (regulated by the EI), a dispatcher sends the corresponding message to AMELI by means of its External Agent.

The *Virtual Worlds Manager* (VWM) mediates all VWs-ECS communications and dynamically updates the 3D representation of all connected VWs by means of aforementioned *Message/Action Dispatchers*. The VWM is composed by one VW proxy for each connected VW. Since different VW platforms can need a different specific programming language, these proxies allow to use such a specific language to communicate with the ECS. In our example scenario we use *OpenMetaverse* (http://openmetaverse.org) library to manipulate the content of *OpenSimulator*.

4.2 Human-agent interactions

As previously introduced, our objective is to facilitate the user a structured hybrid virtual environment for serious purposes (e.g., e-applications). To do so, we provide a VW interface for human participants whilst SW agents are directly connected to the AMELI platform and are represented as bots in the VWs.

We consider three types of interaction mechanisms: *illocution*, *motion*, and *information request*. First, *illocutions* are interactions uttered by participants within activities' protocols. Human avatars interact by means of illocutions by performing gestures and sending chat messages. Bot avatars can do the same except for those representing institutional agents, which can also send public messages by updating information panels. Second, *motions* correspond to movements to enter/exit activities. Human avatars show their intention to (and ask for permission to) enter/exit activities by touching the door of the corresponding room in the VW. As for bots, they are simply teleported between rooms. Third, *information requests* include asking to the institution for information about i) activities reachable ii) activities' protocols states and iii) activities' participants. These interactions have been implemented by both sending messages (e.g. the institution manager sends a private message to an avatar specifying that is not allowed to enter/exit an activity) and drawing on information panels (e.g. the state of an auction is indicated in a panel on a wall of the auction room).

[4]Thus, AMELI perceives all participants as SW agents.

Figure 3: Bot Buyer and human performing bidding gestures in a running auction.

In order to illustrate the communication flow of an interaction between agents and humans, here we describe two communication processes within a negotiation activity. In particular, we detail a bid placement within an auction (see Fig. 3).

The first communication process starts with the desire of a human participant to bid in an auction, so that s/he performs a raising hand gesture with his/her avatar. Then the VWM catches the action and communicates the gesture to the ECS, which uses the *Action Dispatcher* to translate this gesture to the corresponding AMELI message *"bid"*. Afterwards, the *Agent Manager* in the middleware sends such a message to the normative layer. The message is sent by means of the participant's external agent. Next, AMELI processes the message and sends back a response with the result of the message (ok or failure) to the middleware. As a consequence, the middleware uses the VWM to cause (trigger) the action of the market facilitator sending a chat message with the response to all participants within the auction. Notice that, although the bid gesture is always performed by the human avatar, it does not mean that it was a valid action, so the confirmation message sent to the rest of participants is necessary for they to be aware of the action validity.

In the second communication process, a SW agent directly sends a bid message to AMELI, since it is directly connected to the normative layer. Only if the message has been successfully performed in AMELI, it is reflected in the VW. To do so, the middleware receives the said message event from AMELI and translates it by means of the *Message Dispatcher* to the related bot avatar raising its hand. Thus, the human user can perceive bot's bid visually in its VW client. Overall, the human can bid and be aware of all other participants' bid placements. As we have seen, this mechanism allows agents and humans in the same auction activity to interact in a structured and seamless way.

5 EXECUTION EXAMPLE

This section is devoted to briefly describe our *v-mWater* scenario, were human avatars interact with several bot characters. All bots are bold and have differentiated artificial skin colours that represent their roles (see Figs. 3, 4 and 5). Fig. 1 shows three rooms generated in the VW, one for each activity in the EI defined in § 3.2. The institution precinct is delimited by a fence with an entrance on its left side, where the *Institution Manager* restricts the access.

Figure 4: Human seller in the *Registration* room.

Figure 5: The inside of the *Wait&Info* room.

In our example execution, Peter Jackson is a user that controls his human avatar and requests access to the market as a buyer by sending a login message to the *Institution Manager*.

The access to the *Registration* room (see Fig. 4) is limited to participants playing a seller role. There, sellers can register a water right by sending the command *"register ⟨water_right_id⟩ ⟨price⟩"* through the private text chat to the *Market Facilitator* sat at the desktop. Next, the *Market Facilitator* confirms that the registration is valid and sends back the corresponding *"idRegister"* (otherwise, it would send an error message). All correctly registered water rights will be auctioned later on.

All participants are allowed to enter the *Wait&Info* room (see Fig. 5). Several waiting sofas are disposed in this room, a map of the basin is located on its left wall and the desktop located at the end is designated to be used by the *Market Facilitator*. Behind it, one dynamic information panel shows a comprehensive compilation of relevant information about last transactions. The *Market Facilitator* indicates every new transaction updating the information panel. Alternatively, participants can approach the *Market Facilitator* and request for information about last transactions by sending a private chat message *"trans"*. Similarly, they can also request for information about next water rights to negotiate with the private chat message *"nextwr"*. In both cases *Market Facilitator*'s response goes through the same chat.

Buyers can join a negotiation activity by requesting the entrance to the *Auction* room (see Fig. 3). If their access is validated, they can take a sit at one of the free chairs disposed in the room. Two desktops are reserved for the *Market Facilitator* (left) and the *Basin Authority* (right). Then, the bidding process explained before (in § 4.2) takes place. Winner/s will request a desired quantity of water from the *Basin Authority* through a private chat. As a result, the *Basin Authority* notifies the valid agreements to all participants with a gesture and updates the information in the designated panel. Although some details are omitted, we can see that, despite the inherent complexity of the *Auction* activity, it has been designed in a way so that a human participant can easily place bids and intuitively follow the course of a negotiation.

6 USABILITY EVALUATION

This section evaluates our *structured hybrid Virtual Environment* by means of a usability test that follows the widely-used test plan from (Rubin and Chisnell, 2008). First, we define general test objectives and specific research questions that derive from them. Next, we detail test participants and test methodology. Last, we describe and discuss obtained results both at qualitative and quantitative levels.

6.1 Test Objectives

The main goal is to assess the usefulness of our *structured hybrid Virtual Environment*, that is, the degree to which it enables human users to achieve their goals and the user's willingness to use the system. This goal can be subdivided in the following sub-goals: i) assess the *effectiveness* of *v-mWater*, i.e the extent to which users achieve their goals; ii) assess the *efficiency* of *v-mWater*, i.e. the quickness with which the user goals can be accomplished accurately and completely; iii) identify *problems/errors users encounter/make* when immersed on such a structured 3D VE; iv) assess *users' satisfaction*, that is, their opinions, feelings and experiences; and v) open some discussion about the hypothesis that users' *age, gender or skills* may affect effectiveness and user satisfaction.

With all these objectives in mind, we have defined a test task that consists on searching for information about last transactions in the market and registering (for selling) a water right. This structured task is in fact composed of four subtasks: i) *understand the task and figure out the plan* (two out of three rooms have to be visited in a specific order) required to perform the task; ii) *get specific information about the market transactions* at the *Wait&Info* room. This can be accomplished by reading the information panel or rather by asking the Information bot; iii) *work out the required registration price*, which has to be 5€ higher than the price of the most recent transaction; and iv) *register the water right* at the *Registration* room, by talking to the Registration bot.

6.2 Research Questions

With *v-mWater* being a functional prototype, we wanted to answer some questions related to how usable it is, how useful this VE proves to be to different users, and more generally, the users' willingness to perform *e-government* services in VEs. Given the test objectives introduced in the previous section, we address several research questions that derive from them. These questions are divided in two categories. The first category is closely related to the task users are asked to perform in the VE:

RQ1: Information Gathering. How fast does the user find the information needed once s/he enters the *Wait&Info* room? Was the information easy to understand? How did the user obtain that information? (reading a panel or interacting with the agent).

RQ2: Human-bot Interaction. Is the registration desk (and bot) easy to find? How pleasant is the interaction with the bot? Does the user value knowing which characters are bots and which are humans?

RQ3: Task Completion. What obstacles do sellers encounter on the way to the *Registration* room on the VE? What errors do they make when registering a right? How many users did complete the task?

The second category is more general and focuses on user's ability and strategies to move around a 3D virtual space, learnability for novice users, and perceived usefulness and willingness to use VEs for online *e-government* procedures: **RQ4: User Profile Influence.** Does the user profile (age, gender, and experience with computers and VEs) influence perceived task difficulty, user satisfaction and immersiveness?

RQ5: VE Navegability. Which strategy does the user take to move between rooms? Does the user notice (and use) the teleport function? Even noticing it, does s/he prefer to walk around and inspect the 3D space?

RQ6: Applicability to *e-Government*. How do users feel about 3D *e-government* applications after the test? Would they use them in the future?

6.3 Participants

We have recruited 10 participants. They form a diverse user population in terms of features such as age (18-54), gender, computer skills and experience on 3D VEs/games. We find users that have grown with computers and users that have not, therefore we can study how age influences efficiency, perceived easiness, usefulness and their predisposition to use such a 3D and hybrid virtual space for *e-government* related tasks. We also pay special attention to users' computer skills and experience in 3D VEs as it can influence their ability to perform required tasks. Table 1 shows details on participants age, gender, computer skills ('basic', 'medium', 'advanced') and VE/games experience ('none', 'some', 'high'). The classification for computer skills was: 'basic' for participants which use only the most basic functionalities of the computer, such as web browsing, text editing, etc.; 'medium' for users with a minimum knowledge of the computer's internal functioning and who use it in a more complex way such as gaming; and 'advanced' for participants who work professionally with computers, i.e. programmers. Regarding VE skills, 'none' were users who have never used a VE, 'some' described users who have tried it occasionally, and 'high' for users who often use a VE. Notice that although most skills are uniformly distributed, VE experience is strongly biased towards VE newcomers.

6.4 Methodology

The usability study we conducted was mainly exploratory, but somehow summative. We used the Formative Evaluation method (Bowman et al., 2002),

Table 1: List of participants' characteristics.

Name	Age	Gender	PC exp	VE exp
P1	18	Female	Medium	Some
P2	19	Female	Medium	High
P3	20	Male	Advanced	Some
P4	25	Female	Medium	None
P5	25	Female	Medium	None
P6	28	Female	Advanced	None
P7	39	Male	Advanced	None
P8	40	Male	Medium	None
P9	53	Male	Basic	None
P10	54	Female	Basic	None

which fitted our interests at this early iteration of our prototype. We were mostly interested in finding relevant qualitative data. Nevertheless, since the application itself is already a functional prototype, we also took some quantitative measures.

The evaluation team was composed by a moderator and an observer. The former guided the user if needed, encouraged him/her to think-aloud, introduced the test, and gave the user the consent-form and the post-test questionnaire. The latter took notes.

The tests took place at users' locations: half of the participants did the test at their home and the other half at their workplace, on a separate room. The equipment consisted in 2 computers, the VW server and the VW client. The latter recorded user interactions and sound.

All participants were requested to perform a task. Specifically, they were told: *"act as a seller, and register a water right for a price which is 5€ higher than the price of the last transaction done"*. Recall that, in order to do the task properly, participants would then have to visit the *Wait&Info* room, check the price of the last transaction (by asking the bot or checking the information panel), and afterwards head towards the *Registration* room and register a water right at the correct price by interaction with the Registration bot.

The test protocol consists of 4 phases. First, *Pre-test interview*: We welcomed the user, explain test objectives and asked questions about their experience with *e-government*. Second, *Training*: The user played through a demo to learn how to move in 3D and interact with objects and avatars alike. We also showed him the different appearance of bots and humans and gave an explanation of how to interact with bots. This training part was mostly fully guided, except at the end, when the user could freely roam and interact in the demo scenario. Third, *Test*: The user performed the test task without receiving guidance unless s/he ran out of resources. Meanwhile the moderator encouraged the user to think-aloud (by telling him/her to describe actions and thoughts while s/he did the test). Fourth, *Post-test questionnaire*: The user is given a questionnaire with qualitative and quantitative questions regarding *v-mWater* and the application of VEs to *e-government* tasks (see Figure 8).

6.5 Results and Discussion

In this section we discuss usability issues identified after the analysis of data gathered during the test. We will go through the research questions defined in § 6.2. The answers to each of them come from different sources: a combination of the post-test questionnaire; comments given by the users; notes took by

Table 2: Post test questionnaire.

Question Number	Brief description
Q1	Situatedness and movement in 3D
Q2 (Q2.1, Q2.2)	VE walking (2.1) and teleport (2.2) comfortability
Q3	Info gathering (panel/bot)
Q4	Human-bot interaction
Q5	Bot visual distinction
Q6	Chat-based bot communication
Q7	Task easiness
Q8	Immersiveness in 3D
Q9	Improved opinion of 3D VWs
Q10	Likeliness of future usage
Q11	3D interface usefulness
Q12	Overall system opinion
open question	User's comments

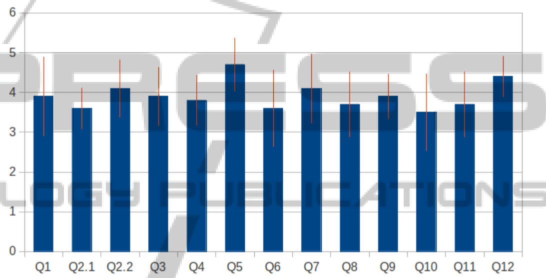

Figure 6: Post-test questionnaire results. X axis: questions from Table 2. Y axis: average (and standard deviation) values.

the observer; and the review of the desktop and voice recordings that were taken during the test (i.e. while participants performed the task).

Table 2 summarizes the 12 questions in the post-test questionnaire, and Figure 6 depicts a compilation of users' answers. There, X axis shows each of the post-test questions and the Y axis shows average values of answers considering a five-point Likert scale. This scale provides 5 different alternatives in terms of application successfulness ('very bad'/'bad'/'fair'/'good'/'very good'), where 'very bad' corresponds to 1, and 'very good' to 5. Standard deviation values are also provided.

Overall, the quantitative results we obtained from the questionnaires were very satisfactory, with all average answers higher than 3.5 (standard deviation lower than 1.0). Highest rated responses (whose values were higher than 4.0) were associated with the easy distinction of bots and human controlled characters (Q5) and the overall satisfaction of the user (Q12). On the other end, lowest rated responses (with 3.5 values) were related to the comfortability when walking within the environment (Q2.1), the command system used to chat with the bots (Q6), and the idea of using a 3D VE for similar tasks (Q10).

From both the qualitative measures that the user gave at the open question of the post-test questionnaire as well as when debriefing with the evaluating team, we extracted a number of relevant aspects of the *v-mWater* system. Firstly, users like its learnability, its immersiveness, and how scenario settings facilitate task accomplishment. Moreover, users like 3D visualization although as of today, it is too soon for them to imagine a VE being used for everyday tasks, since it is hard to imagine, unfamiliar, and in some cases users wouldn't fully trust on it. At the same time, the overall opinion of the system was positive and some users clarified that they were not entirely comfort using the application, but they would easily become used to it; since it was highly learnable and safe to use.

Usability criteria, such as effectiveness, efficiency and errors have been analysed answering the research questions from first category introduced in § 6.2.

RQ1: Information Gathering. The information that the user had to obtain in subtask ii) could be gathered from 2 sources: the information panel and the Information bot, both located at the *Wait&Info* room. During the test, the majority of the users, except two of them who did not enter this room, walked directly towards the information panel and/or the information desk (where the bot was located). These users could easily read the information from both sources. Answers of Q1 and Q3, both with an average close to 4, reinforce previous statement.

RQ2: Human-bot Interaction. Users should interact with bots in subtasks ii) and iv). The high average of Q4 indicates that the user had a good overall impression about human-bot interaction. Nevertheless, Q6 denotes that users were uncomfortable with the technique, a command-based system, used during the dialogue with the bot. Analysing Q5, with an average of 4.7, we can state that participants found it almost imperative to know when they were facing a bot.

RQ3: Task Completion. Overall, participants found it easy to complete the task (as Q7 indicates with an average of 4), and they took an average of 4.46 minutes. Users have not found any obstacles that prevented them from completing the task. Regarding errors that users committed during the task completion, some users did not always go to the right destination (building), but they always realised their mistake and were able to get to the correct destination. Another type of error relates to the chat-based interaction with bots; as Q6 indicates, where the average of the answers was 3.6. Users with low computer skills had some trouble when interacting with the bot because of the strict command-based system. Nevertheless, the users found this communication system highly learnable. Related to the effectiveness of the application, we have measured it re-viewing the desktop recordings. Considering the structure of the task that has been detailed in § 6.1, the percentage of users that completed the corresponding sub-tasks were: i) 80% understood the task correctly. Only 20% of users did not figure out they had to check prices before registering their water right. ii) 80% of users gathered the information correctly (the rest skipped that step). iii) 70% of users calculated the price properly. iv) 100% completed the registration subtask, i.e. all participants registered water rights.

Below, we give a brief discussion about user profile influence on perceived task difficulty, satisfaction, usefulness and immersiveness, and analyse more general usability aspects of our system such as the user's ability to move around a 3D VE; or perceived usefulness of VEs for on-line *e-government* procedures.

RQ4: User Profile Influence. This question was answered by analysing the results from our post-test questionnaire in terms of user features. From the point of view of age, participants are equally balanced. As the age increases it also does the difficulty to use the application, although the satisfaction also increases. Surprisingly, the youngest users found the application less useful than the older ones (this may be due to their higher expectations from 3D VE). Related to users' experience with computers, users with the lowest experience had clearly a harder time using the arrow controls to walk around the 3D space. Additionally, this group found difficult both the interaction with the bots and the task completion. Similarly, the immersion grows as the experience with computer grows.

RQ5: VE Navigability. Navigation in our VE has proven to be relatively easy, since users' average opinion was 4 (Q1). They did not roam in any occasion as it has been appreciated on the recordings. Users who found out they could teleport were comfortable using it, as they reflected on the post-test questionnaire (Q2.2) and also by some of their comments.

RQ6: Applicability to *e-Government*. Users' opinion about VEs had improved after doing the test (Q9), since they answered with an average value of 4. When asked about their intention to use a similar system for similar tasks in the future (Q10), users answered an average of 3.5, which means that they have a relative good opinion about the usefulness of the application. Finally, users reported that the 3D interface had helped them in achieving their goals during the test, as Q11 shows with an average value of 4.

Finally, we were able to extend the test with 3 additional PC and VE expert users. Since they were comfortable with the controls, they completed the task notably faster than newcomers. Moreover, all answers got higher marks on average. This assesses the potential of this approach. Nevertheless, the command-based human-agent interaction still appeared as the weakest feature.

7 CONCLUSIONS

In this paper we have explicitly structured participants' interactions in hybrid (humans and software agents) virtual environments (VE). We have presented an example scenario in an e-government application (*v-mWater*, a *v*irtual *m*arket based on trading *Water*), and evaluated its usability. We have also described the execution infrastructure that supports this hybrid and structured scenario where humans and bots interact both each other and with the environment. Furthermore, we characterize different interaction mechanisms and provide human users with multi-modal (visual, gestural and textual) interaction. In our usability study, we have paid special attention to how users perceive their interaction with bots.

The usability evaluation results provide an early feedback on the implemented scenario. *v-mWater* is perceived as a useful and powerful application that could facilitate everyday tasks in the future. Users like its learnability, its immersiveness, and how scenario settings facilitate task accomplishment. In general, users have well completed the proposed task and were able to go to the right destination in the scenario. After doing the test, users improved their opinion about 3D VEs. In addition, the overall opinion of the human-bot interaction is positive.

Nevertheless, there are some inherent limitations of interface dialogs and interactions. Some users are not comfortable using the command-based bot dialog and find difficult to move their avatar in the 3D VE. Thus, a future research direction is to define new forms of human-bot interactions, using multimodal techniques based on voice, or sounds and tactile feedback supported by gaming devices. We also plan to incorporate assistant agents to help humans participate effectively in the system, and perform a comparative usability study to assess assistants' utility.

REFERENCES

Almajano, P., Trescak, T., Esteva, M., Rodriguez, I., and Lopez-Sanchez, M. (2012). v-mWater: a 3D Virtual Market for Water Rights. In *AAMAS '12*, pages 1483–1484.

Behrens, T., Hindriks, K., and Dix, J. (2011). Towards an environment interface standard for agent platforms. *Annals of Mathematics and Artificial Intelligence*, 61:261–295.

Book, B. (2004). Moving beyond the game: social virtual worlds. *State of Play*, 2:6–8.

Bowman, D., Gabbard, J., and Hix, D. (2002). A survey of usability evaluation in virtual environments: classification and comparison of methods. *Presence: Teleoperators & Virtual Environments*, 11(4):404–424.

Campos Miralles, J., Esteva, M., López Sánchez, M., Morales Matamoros, J., and Salamó Llorente, M. (2011). Organisational adaptation of multi-agent systems in a peer-to-peer scenario. *Computing, 2011, vol. 91, p. 169-215*.

Cranefield, S. and Li, G. (2009). Monitoring social expectations in Second Life. In *AAMAS'09*, pages 1303–1304, Richland, SC.

Dignum, V., Meyer, J., Weigand, H., and Dignum, F. (2002). An organization-oriented model for agent societies. In *AAMAS'02*.

Esteva, M., Rosell, B., Rodríguez-Aguilar, J. A., and Arcos, J. L. (2004). AMELI: An agent-based middleware for electronic institutions. In *AAMAS'04*, pages 236–243.

Ferraris, C. and Martel, C. (2000). Regulation in groupware: the example of a collaborative drawing tool for young children. In *CRIWG'00*, pages 119–127. IEEE.

Giret, A., Garrido, A., Gimeno, J. A., Botti, V., and Noriega, P. (2011). A MAS decision support tool for water-right markets. In *AAMAS '11*, pages 1305–1306.

Guerra, A., Paredes, H., Fonseca, B., and Martins, F. (2008). Towards a Virtual Environment for Regulated Interaction Using the Social Theatres Model. *Groupware: Design, Implementation, and Use*, pages 164–170.

Mezura-Godoy, C. and Talbot, S. (2001). Towards social regulation in computer-supported collaborative work. In *CRIWG'01*, pages 84–89. IEEE.

Paredes, H. and Martins, F. (2010). Social interaction regulation in virtual web environments using the Social Theatres model. In *Computer Supported Cooperative Work in Design (CSCWD), 2010 14th International Conference on*, pages 772–777. IEEE.

Pronost, N., Sandholm, A., and Thalmann, D. (2011). A visualization framework for the analysis of neuromuscular simulations. *The Visual Computer*, 27(2):109–119.

Rubin, J. and Chisnell, D. (2008). *Handbook of usability testing : how to plan, design, and conduct effective tests*. Wiley Publ., Indianapolis, Ind.

Trescak, T., Esteva, M., and Rodriguez, I. (2011). VIXEE an Innovative Communication Infrastructure for Virtual Institutions (Extended Abstract). In *AAMAS '11*, pages 1131–1132.

Tromp, J., Steed, A., and Wilson, J. (2003). Systematic usability evaluation and design issues for collaborative virtual environments. *Presence: Teleoperators & Virtual Environments*, 12(3):241–267.

Tsiatsos, T., Andreas, K., and Pomportsis, A. (2010). Evaluation framework for collaborative educational virtual environments. *Educational Technology & Society*, 13(2):65–77.

van Oijen, J., Vanhée, L., and Dignum, F. (2011). CIGA: A Middleware for Intelligent Agents in Virtual Environments. In *AEGS 2011*.