Journal of Universal Computer Science, vol. 18, no. 18 (2012), 2556-2575
submitted: 31/5/12, accepted: 28/9/12, appeared: 1/10/12 © J.UCS

Creating Test Questions for 3D Collaborative Virtual Worlds: the WorldOfQuestions Authoring Environment

María Blanca Ibáñez
(Universidad Carlos III de Madrid, Leganés, Spain
mbibanez@it.uc3m.es)

José Jesús García Rueda
(Universidad Carlos III de Madrid, Leganés, Spain
rueda@it.uc3m.es)

Diego Morillo
(Universidad Carlos III de Madrid, Leganés, Spain
dmorillo@it.uc3m.es)

Carlos Delgado Kloos
(Universidad Carlos III de Madrid, Leganés, Spain
cdk@it.uc3m.es)

Abstract: In this paper we introduce, describe and evaluate WorldOfQuestions, an authoring environment for the creation and delivery of tests in 3D Collaborative Virtual Worlds. This environment is composed of an extended, customized version of the Open Wonderland platform and a form-based editor. Its aim is to make the most of 3D world features, such as immersion and interactivity, when implementing multiple choice, ordering and essay questions enriched with multimedia elements and 3D objects. A group of teaching professionals was asked to work with the environment, in order to evaluate its usefulness and ease of use. A Technology Acceptance Model (TAM) based framework was used for this evaluation. According to the educators interviewed, the most important aspects to consider regarding the behavioral intention to use the tool were the academic subject being taught, the student profiles, the environment learning curve and the time requirements.

Keywords: Authoring tools, 3D Collaborative Virtual Worlds, Computer Based Assessment, TAM
Categories: L.3.0

1 Introduction

For quite a number of years now, information and communications technology (ICT) has been permeating the educational world in all its aspects, from information access to group-work and lab practices, and has also been embracing a varied range of pedagogical approaches. Activities of such importance as assessing student performance and outcomes have also been influenced by Computer Based Assessment (CBA) which brings a promise of mechanization of the assessment process, reuse, customization, security, lower cost and automatic record keeping ([Chatzopoulou 10], [Smith 05]).

In addition, educators and educational researchers have recently taken an interest in 3D Collaborative Virtual Worlds (3DCVWs) such as Second Life [Minocha 10], OpenSim [Delp 07] and Open Wonderland [Gardner 11], highlighting their immersive and realistic capabilities as a way to foster student motivation and engagement ([Chittaro 07], [Dalgarno 10]). 3DCVWs provide the illusion of a 3D space where people, as avatars, can interact with the 3D objects populating the environment, and communicate with one other ([Bainbridge 07], [Davis 09]). However, until now research studying the assessment processes in 3DCVWs seems quite scarce: even when a virtual world is used as part of an educational activity, assessment typically relies on more traditional, non-CBA mechanisms or, if using CBA, on Web2D resources; such as Sloodle [Livingstone 08].

In this paper, we describe WorldOfQuestions, an authoring environment for the creation, deployment and delivery of tests in 3DCVWs. The aim of this tool is to make the most of the characteristics of these virtual worlds in order to produce enriched tests, but at the same time make it easier for teachers to define questions for this 3D environment, and to link them to elements in the virtual world. Developing educational content for 3DCVWs is a time-consuming and complex activity, and the creation of tests is no exception. With this project, we aim to make teaching professionals more willing to embrace this technology and to use it for assessment purposes by making the creation process more enjoyable, less time-consuming and easier. Teachers and professors can focus on the advantages offered by virtual worlds because the drawbacks have been reduced. The building and application of WorldOfQuestions seeks to contribute to a better understanding of an emerging technology in a critical aspect of learning by following the general principles of design-science research ([Hevner, 2004]).

The remaining sections of the paper are organized as follows. We first briefly summarize the interactive capabilities of 3DCVWs and their uses in assessment. We then describe the WorldOfQuestions tool as well as the authoring process it promotes. Next, we describe our research methods and present the results of the evaluation of the authoring environment by a group of education professionals. Finally, we conclude the paper with a discussion of these evaluation results.

2 Interactive capabilities of 3D Collaborative Virtual Worlds. Opportunities for assessment

3D Collaborative Virtual Worlds can be seen as a multimodal user interface where 3D scenarios are filled with interactive 3D objects, which can be either the representation of real or fictional objects, and of synthetic characters able to act plausibly [Bell 08], [Calogne 08]. Additionally, 3DCVWs allow truly immersive spaces to be deployed, thus fostering the learner's imagination and offering rich possibilities for interaction, through avatars, with the environment, objects and other community members [Dillenborg 02], [Eschenbrenner 08], [Girvan 10], [Kallonis 10]. In education, interaction with spaces and objects is useful for training and simulation, whereas interaction with partners is useful for collaborative activities and awareness [Dickey 05], [Dalgarno 10].

Nowadays, 3DCVWs are considered to offer a brand-new, exciting place for learning and teaching and authors have issued guidelines for using different learning strategies, such as situated learning, role playing, cooperative-collaborative learning and problem-based learning, in 3DCVWs [Dillenbourg 02], [Chittaro 07], [Huang 10]. The stated principles suggest the use of visual elements of 3DCVWs to immerse students in a situation where the problem to be solved or the concepts to be learnt are presented in a contextualized and natural way. Furthermore, 3D objects and synthetic characters can be used not only as instruments to transmit information, but also as tools to build knowledge following constructivist principles [Ibáñez 10]. However, despite the relevance that assessment has in the learning process [Bradsford 00], there are an insufficient number of 3D collaborative learning environments that incorporate assessment activities [Dalgarno 10], [Mikropoulos 11], [Savin-Baden 08].

Assessment in 3DCVWs follows two different approaches, namely automated versions of item-based paper-and-pencil tests and performance-based assessment. The first promises to be more cost effective and easier to scale whereas the second is based on richer observations about student learning [Clark 10]. Performance-based assessment has been used in 3D video games where some studies highlight the importance of not breaking the flow state, induced by immersive games, with invasive assessment activities [Shute 09], [Kickmeier-Rust 08]. It has also been used in V. J. Shute's work [Shute 09] through Bayesian models to monitor actions, integrate evidence on learner performance, and update the student model in relation to competencies. And the most representative work that follows this approach is probably River City [Nelson 10], a multi-user virtual environment, where detailed records of students' actions are collected and stored to assess how learners detect and decipher illness patterns in a town besieged by health problems. QuizHUD [Bloomfield 09] represents the first attempt to integrate a computer assisted assessment tool into a 3D virtual world. QuizHUD includes both classical multiple choice and exploration questions. Classic multiple choice questions are included in the virtual world as they would be in a web-based learning environment: using 2D panels. For the exploration questions, students must click on 3D objects to state their choice.

Our work follows the exploration question approach and extends it by using a richer set of interactions within the 3D learning environment and minimizing the intrusion of 2D assets in the 3D interface. The ultimate aim is to provide learners with immersive assessment activities that are naturally integrated in collaborative learning experiences in 3D virtual worlds.

2.1 Interaction of users with a 3D assessment system

Typical assessment tests are composed of a set of questions with a predetermined correct answer, a grading scheme and possible feedback. A typical closed-choice test question consists of the following parts:

1. Wording to convey the question to students.
2. A response template with the list of choices.
3. A correct response.
4. Feedback and grades for different options.

When dealing with computer assisted assessment, the system will present the wording to the students, who will choose an answer from among the choices provided by the system, and subsequently the system will offer feedback on the accuracy of the selected answer.

In a 3DCVW, students immersed in the environment perceive and act through their avatars. Perception comes from the senses of hearing and sight and can be quite realistic. Although it is technologically feasible, haptic perception and action are not typically included in 3D virtual environments.

3DCVWs can convey visual information to users through text, pictures, audio and video fragments, 3D objects and the 3D setting where the learning activity is deployed. Audio information is usually transmitted in combination with visual information through video or actions carried out by 3D objects (including synthetic characters).

Wordings can be made available to users in different ways:

- Text, as is usually done in a classroom or a web-based learning environment.
- Audio or video file.
- Spoken words and actions by one or more synthetic characters.

It is also possible to adapt the system to student preferences for receiving information. Additionally, mechanisms should be provided to replay the information as many times as necessary for full understanding.

Once students have received a test question, a list of choices is available for them. This list can be in the shape of, for instance:

- 3D objects, including synthetic characters, within the virtual world.
- A set of different places in the virtual world.

Feedback can be received in the same way as wording. Feedback may also involve changing the 3D scenario or any of its elements. For instance, a correct answer to a question may cause a song to start playing or a 3D object that represents a prize to appear or the student's avatar to be teleported to another place in the virtual world, or the occurrence of all these actions simultaneously.

On the other hand, avatars can perform a rich variety of actions that can be monitored by the system and thus can be interpreted as answering a question:

- Movement from one place to another either by walking, running or teleporting. The information to be considered as an answer can be either the destination point, the proximity to an object or the path followed. The time used can also be recorded.
- Making gestures to indicate yes or no, waving or clapping, to name just a few. Gestures represent a way of communication that has not yet been employed in traditional test-based assessment.
- Including assets in the virtual world: images, videos, audio-recordings.
- Manipulation of 3D objects including the selection or movement of an object, modification of its size or its use in any kind of simulation.

- Interaction with synthetic characters. Synthetic characters can have simple conversations with students thanks to the Artificial Intelligence techniques currently used by chat bots.

3 The assessment authoring environment: WorldOfQuestions

WorldOfQuestions (WoQ) is an assessment authoring tool that facilitates the process of creating tests to be delivered and enacted in 3DCVWs. The tool embeds the WoQ-editor, the WoQ-architecture and the 3D collaborative virtual world platform Open Wonderland. Tests are created using Open Wonderland together with the WoQ-editor, and are then delivered as 3D collaborative learning environments by WoQ. Once a test created by the WoQ-editor becomes a 3D collaborative learning environment, it can be executed by the WoQ-architecture as many times as desired.

3.1 Open Wonderland: a multi-user platform to support WoQ

Tests created using WoQ are delivered as 3D collaborative learning environments within a multi-user 3D virtual world platform: Open Wonderland. Open Wonderland facilitates the test creation process by allowing the inclusion of multimedia assets from different sources via drag-and-drop. For instance, 3D objects exported from Google 3D Warehouse and avatars and synthetic characters from Evolver can be used; pictures in jpg image file format from the desktop; as well as videos in a variety of formats including wmv, mpeg4, and avi also from the desktop.

Open Wonderland offers particular features that are useful to support an 3D virtual world assessment engine. It has capabilities to assign users to groups along with a security mechanism that restricts interactions with assets. The combination of these capabilities opened up the possibility to create collaborative assessment activities that can be visible only to subsets of students. Moreover, restricted interactions were used to maintain the workflow of tasks associated with a question, namely wording, answers, correct answer and feedback. Finally, Open Wonderland provides scripting mechanisms for performing actions as responses to events on 3D assets; this characteristic was useful in building a live assessment learning environment.

Open Simulator and Open Cobalt, among others, are 3D virtual world open source platforms that are also suitable for our purposes.

3.2 Question types supported by WoQ

WoQ offers a reduced number of available question types selected from common lists of test question types, taking into particular account the types of questions collected in Question & Test Interoperability (QTI) [IMS 06], according to three criteria: relevance, adaptation and extensibility.

The selected question types had to be relevant, i.e. widely used in assessment and easily adaptable to any area of knowledge. *Multiple choice*, for instance, is probably the most commonly used test question type, so it had to be included in our system. In a 3D world, these answers can take the shape of 3D objects or characters, multimedia fragments, areas in the world, etc., and users can select one of them by clicking on it,

coming closer to it or entering a certain building or area. As teachers often need to complement multiple choice based assessment with more subjective input from their students, *essay* type questions had to be included also. By using this type of question teachers have the chance to ask their students to write a short text to answer an open-ended question. In a 3D world any element suitable for writing, such as blackboards or panels, can be used as a means to submit an answer. For this kind of question, automatic correction is not possible.

The question types offered by our system had to be well adapted to the very nature of 3D virtual worlds, i.e. they had to have an especially interesting role to play in this kind of environment. *Putting things in order*, apart from being quite relevant from an assessment point of view provides an opportunity to manipulate 3D content, to make students reason on the very spatial distribution of the 3D world, to move things around and change them and compare elements in the world. The range of elements to be put into order in a 3D world is very wide: you can put certain 3D objects or characters, areas in the world, etc, into order.

As an example, two test questions about the Football World Championship will be fully described, illustrating how the interactive capabilities of 3DCVWs and WoQ can be used.

The wording of the first question ("Who won the last Football World Championship?") is delivered by a text message, when students click on a 3D object representing a trophy. At the same time, a set of synthetic characters appear in the world next to the student's avatar. Each character is wearing the uniform of a national football team (Spain, Italy and Brazil). Students have to provide an answer by clicking on one of these characters. The feedback to the answer is a video displayed in-world. The entire sequence can be seen in Figure 1.

The wording of the second question ("Select the scoreboard showing the final goal of the last match.") is delivered as a text, reinforced with an audio message, when students get closer to a 3D object representing a ball. This message involves additional objects that appear in the world: three scoreboards showing the pictures of three different shots. Students will have to answer the question by having their avatars go to the place in the world (a football stadium) where the scoreboards are. The feedback to their answer is a video displaying the actual goal.

These examples show how 3D and multimedia features can be added to test questions in order to make them richer. It is expected that, by using these kinds of features, such as moving around, going to places or searching for locations, interacting with 3D objects, watching videos, etc., the students' feeling of immersion will increase, and thus their involvement in assessment tasks will improve. Participants in these learning experiences can even modify the world as a consequence of a question, in that way the setting for the next question is modified as desired, and that allows for a higher degree of test customization, as the new objects cannot be seen by all participants, only by the ones who answered the question in a specific way.

3.3 Authoring process with WoQ

The authoring process involves the use of a 3D virtual world delivery platform (Open Wonderland) and the WoQ-editor. Authors start within the 3D virtual world delivery platform and modify the 3D setting by adding all the assets that will be part of the

questions being created: 3D objects, synthetic characters, videos, audios, pictures, etc (see figure 2).

Once all the necessary assets have been properly placed in the world, the next step in the authoring process involves the creation of a new snapshot. A snapshot is a copy of the current state of the virtual world. That copy will be used later as the setting for the delivery of the questions.

Figure 1: Deployment of the question ("Who won the last Football World Championship?")

The next step is to use the WoQ-editor (see figure 3) to define each question, specifying its wording, possible answers, correct answer and feedback, and to relate all these fields to the assets previously added to the virtual world. The editor allows the inclusion of *hints*, which can be used to help students to identify the places and objects in the 3D world where wordings are provided.

Figure 2: Editing an asset's properties in Open Wonderland.

The WoQ-editor has all the necessary functionality to deal with visibility issues: in Open Wonderland, not all questions or their related assets are visible for all students at all times. By using the visibility features included in WoQ, any asset

related to a question (triggering object, answers and feedback), and the question itself, can be visible for some students, but not for others.

Once all the questions have been created, they can be grouped to form a complete test.

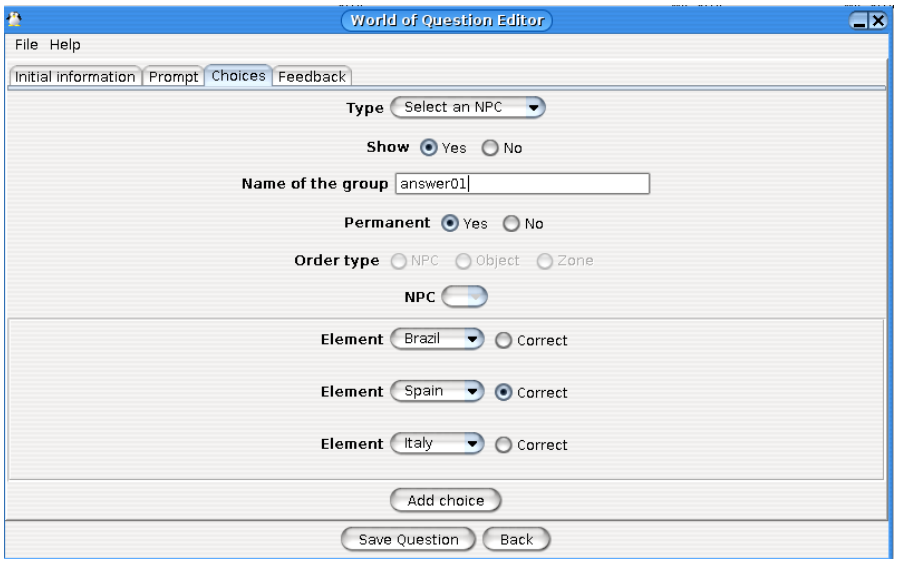

Figure 3: WorldOfQuestions editor.

3.4 Architecture of WoQ.

The architecture of the system is shown in figure 4. This architecture is built on a previous extension of the Open Wonderland platform's client-server architecture, made by our team for an earlier project.

Open Wonderland's extended server has now been complemented with two additional repositories: a first repository storing educational information (the questions files, in WoQ format, and users' scores) and a second repository storing all assets involved in the assessment activity (multimedia files and 3D assets).

Two processing modules have been added to Open Wonderland's extended client: one module is used to actually orchestrate the sequence of questions, answers and feedback, while the other one is entrusted with monitoring user actions and dynamically managing the visibility of each asset for each user or group of users. The orchestrating module has been divided into two processing levels. The High Level Orchestration Module decides on the next question to be delivered to the user, while the Low Level Orchestration Module actually delivers the selected question, waits for an answer and offers feedback. The user interaction managing module is also organized into two different sections, separating monitoring tasks (i.e., listening to user actions) from visibility managing tasks.

Once the assessment engine is running, the HLOM will ask the Assessment Repository for a list of all questions in the repository. When a question is activated, the HLOM sends it to the LLOM for delivery and display of the question's wording.

In order to fulfil this task, the LLOM will ask the Assessment Repository for the question's wording and available answers, and will indicate to the Visibility Manager what assets are to be made visible and to what users. The Visibility Manager will give users the necessary permissions to view the pertinent assets, which have been obtained from the Assets Repository, and will be displayed in the virtual world by the Front End.

Once the user chooses an answer (by for instance walking into a specific area, or clicking on an object or approaching a synthetic character) the Monitor will transmit the selection to the LLOM. The LLOM will check the accuracy of the answer by consulting the Assessment Repository, and will decide on the feedback to be shown to the user –also obtained from the Assessment Repository. At the end of the process the Assessment Repository will receive the user's score, as well. The appropriate feedback will be sent to the Visibility Manager for display.

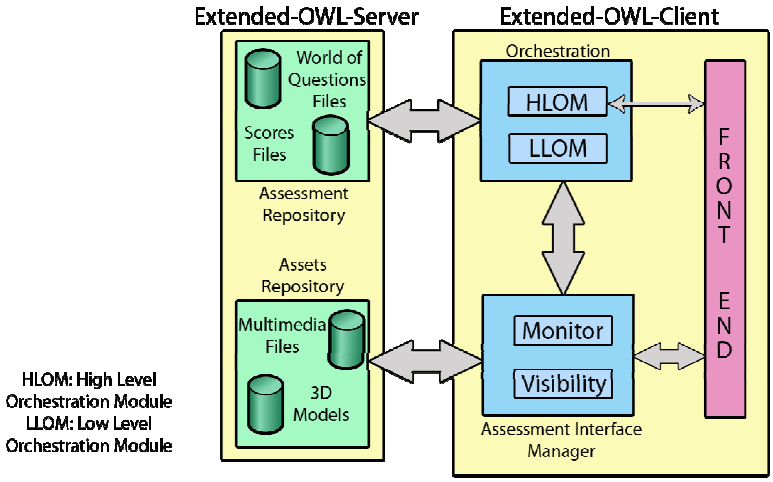

Figure 4: System architecture.

After completion of this last step, the HLOM will update the list of available questions and the system will wait for the user to activate another question.

4 Evaluation of WorldOfQuestions

In order to evaluate the usefulness of WoQ, an experiment was set up in which six experienced university professors and one high school teacher had the opportunity to work with the authoring environment. They then reported on the difficulties they encountered, their willingness to use it in their assessment activities and the motivating effects WoQ could have on both students and teachers.

The participants were invited to create a two-question test using the WoQ assessment tool and were provided with the assets required to create the two pre-defined questions. The participants performed the following tasks:

1. *Modification of a 3D virtual world in Open Wonderland*. This activity included the following tasks: drag-and-drop of the provided assets into the 3D setting; moving them to the right locations; changing their names and appearances and, finally, creating a snapshot.
2. *Creation of the test using the WoQ-editor*. On the previously created snapshot, participants tied assets to each part of the questions using the WoQ-editor interface. Once the two required questions were defined, the participants had to embed them into a test.
3. *Deployment of the test in Open Wonderland*. This activity consisted of the final configuration of the 3D collaborative virtual learning environment and the enactment of the test.

Even though the final objective of the experiment was to gather as much information as possible on users' opinions, expectations and feedback regarding WoQ, we organized all the elements in the experience to revolve around one main research question:

- According to the educators, what are the main factors that influence the intention of use of 3DCVWs in general, and WoQ in particular, for assessment?

The expectation was for this question to be highly influenced by these other two:

- Do teachers and professors feel confident of their ability to use 3DCVWs in general, and WoQ in particular, for assessment purposes?
- How positive do they consider the added value in the use of 3DCVWs in general, and WoQ in particular, for assessment? Would they use them for summative evaluation?

4.1 Measurements

As our measurement frame, we chose Technology Acceptance Model 3 (TAM3) [Venkatesh 08], an information system theory that models how users come to accept and use a technology, based on the relationships between usefulness, ease of use and system use. TAM is probably the most frequently used acceptance model today, and version 3 is the most up-to-date. It covers all the aspects required by this study, while limiting the number of elements to be taken into account in the evaluation.

Following TAM's principles, participants in the experience were asked several questions related to the measurement dimensions in TAM:

- Perceived Usefulness: Would the use of WoQ improve your assessment activities and lead you to obtain better measures of your students' performance? Would the use of WoQ increase your students' motivation, to the point of willingly facing more challenging tests?
- Perceived Ease of Use: Would it be easy for you to become proficient at using the system? Is it a serious problem to work simultaneously with a 3D virtual world and a form-based editor?

- Facilitating Conditions: Is the information provided by the tool and the tutorial enough to learn how to use WoQ?
- Perceived Playfulness: Would the use of WoQ make the task of preparing exams more enjoyable? Would it stimulate your curiosity?
- Behavioral Intention to use the system: Would you use WoQ for your assessment activities? Would you use it for summative evaluation, or even for a final exam?

We included two additional dimensions, as suggested in [Terzis 11]:

- Content: Are the types of questions available in WoQ useful, sufficient, and relevant? Are the alternatives for expressing wordings, answers and feedback sufficiently flexible and enriching?
- Goal Expectancy: How do you feel about using WoQ for assessment purposes in the future? Would you find it too troublesome to come up with new questions to be implemented with the environment?

4.2 Procedure

Each of the participants was given an appointment to work individually with the authoring environment. Two researchers were present at all times: one for technical support and the other to observe the participant's behaviour, voiced questions and general interaction with the system. Each experience was planned to last ninety minutes.

Participants answered an exploration questionnaire before starting their work. The questionnaire included demographic questions on the participants.

After a short introduction to 3D virtual worlds and the environment itself, they were given a tutorial with the final objective of creating two questions.

The tutorial was a detailed working guide of WoQ, including basic operation with Open Wonderland, a full explanation of every menu and entry in the WoQ-editor, configuration of certain files and allocation of files (multimedia files, snapshot and XML files) in the proper sub-directories in the computer's file system.

All the needed assets (multimedia files and 3D objects), including the 3D setting, were provided in advance.

The last part of the experience consisted of the final configuration of the 3D world and test delivery.

Participant working time on the different activities was measured by one of the researchers.

Once the experience was completed, each participant was interviewed following a pre-structured questionnaire that included the parameters listed in the previous section.

5 Results

In this section, we will proceed to show the results obtained during the experience, regarding the different measures in section 4.1. But first we will briefly review each participant's working time, and its relationship with their previous experience.

5.1 Working times and previous experience.

Each participant's working time during the experiment and previous experience with 3DCVWs are summarized in table 1.

In general, working with the 3DCVW took longer than working with the Editor, as expected, although there were notable differences among users: users carry out their activities in the virtual world most quickly were those with previous experience in virtual worlds, while inexperienced users had to spend more time allocating the different assets in the 3D setting. It was an especially tricky task for the high school teacher (participant 7) and one of the university professors who stated having never had any interest in video games or other forms of interactive digital entertainment (participant 1).

Participant	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Previous Experience	*No*	*User and Teacher*	*No*	*No*	*User*	*Student and Teacher*	*No*
Time working with the 3DCVW (min)	*60*	*15*	*35*	*35*	*20*	*35*	*60*
Time configuring and copying files (min)	*30*	*30*	*25*	*10*	*20*	*20*	*25*
Time working with the Editor (min)	*30*	*15*	*25*	*30*	*20*	*35*	*25*
Total working time (min)	*120*	*60*	*85*	*75*	*60*	*95*	*110*

Table 1: Working times and previous experience.

However, moving files between directories in the file system and changing some configuration data, a task that a priori was expected to take just a few minutes, took longer. Most participants expressed that the reason for this was they were not familiar with the directory structure in the computer used during the experience, and that it would have been different if they had been using their own computers.

5.2 Perceived Usefulness

Almost all of the participants (5 out of 7) agreed that this environment would let them design better contextualized tests. Furthermore, they were convinced that tests created and delivered with this environment would improve their students' motivation, although some had the opinion that student profile and subject being taught were, again, fundamental factors.

On the other hand, even if they were inclined to have a positive perception, there were different opinions among the participants regarding whether the use of WoQ would improve their assessment activities, and whether it would allow them to obtain better measures of their students' performance or create more challenging tests (4 participants gave it a positive evaluation). For most of them it depended strongly on the specific subject to be assessed and on student profile. In any case, they agreed that it would enrich the information obtained during the assessment process, as complementary knowledge. Regarding this, one of the participants stated:

"The use of WoQ would allow me not to assess better, but to assess differently".

5.3 Perceived Ease of Use

It is worth distinguishing between the perceived ease of use of the 3D virtual world platform and the WoQ-editor itself. In the first case, participants reported quite a number of problems to spatially place objects in the virtual world, and configuration of synthetic characters caused some confusion. In Open Wonderland, some changes in character configuration are not immediately visible, and most participants complained about this. On the other hand, WoQ-editor's usability was very well received by all of the participants. Defining each element in the questions (hint, wording, answers and feedback) and linking the elements to 3D assets in the virtual world was considered straightforward and easy.

The most important drawback in the authoring process was closely related to groups and visibility: all of the participants found it difficult to deal with visibility and group management. They could not understand how the groups worked and how to configure them.

"Managing groups is certainly excruciating!" said one of the participants.

Finally, all participants agreed that they could easily master the tool, once the conceptual leap was made and the main concepts were properly understood, and that remembering how to operate the system would not be difficult either. Working with a 3D virtual world and a form-based application at the same time was not perceived as a problem at all. In fact, some participants remarked that the two environments complemented each other very well, and that integrating them into one single interface would be nice but not necessary, even if by doing so there would be no need to move files across the file system, which could decrease implementation time:

"A regular user might prefer a unified interface, but it is not a problem to keep them separate".

Finally, almost all of them (5 out of 7) were confident that they would be able to bring their ideas about new test questions to life using WoQ.

5.4 Facilitating Conditions

All of the participants agreed that the information provided by the system and the guiding tutorial was enough to learn how to use WoQ, and also sufficient to use the tool on their own. Most of them would enthusiastically embrace the task of creating new questions with WoQ. Even so, some users suggested the addition of a consulting reference:

"The tutorial is fine, but for later use a quick reference would be lovely".

For 5 of the participants, the WoQ-editor was easy to learn because it is remarkably similar to other editors.

5.5 Content

Participants agreed that the types of questions offered by the system were both useful and sufficient, from a general point of view. When asked to be more specific, some expressed the opinion that the true usefulness of those types of questions would be more apparent in courses with a strong audiovisual bias, and that in the future it would be nice to include questions that involved moving objects or uploading files:

"If you want more types of questions, what about questions involving the movement of objects?"

The same shared positive reaction was expressed regarding flexibility of the mechanisms to provide an answer to the questions, and the richness provided by the different ways of giving students feedback: *"Using both you can take different types of learners into account when creating a test"*. Also some participants suggested including the score in the feedback, and more explicitly using gamification.

5.6 Goal Expectancy – Satisfaction.

Even if there was a tendency among participants to think positively about their willingness to use WoQ in the future, 3 participants expressed some doubts about their capability, as educators, to conceive of new questions that could be implemented and delivered using WoQ. They were again concerned about the necessary conceptual leap and about the usefulness being dependent on the academic subject and area of knowledge.

"To be honest, I can't stop thinking that this could involve an overly profound conceptual change for many teachers..."

5.7 Perceived Playfulness.

All participants agreed that using WoQ would foster teachers' curiosity and willingness to learn more about the possibilities that 3D environments offer as

teaching devices, as they would need to be creative in order to bring engaging, effective and motivating questions to life:

"No doubt you have to use your creativity in order to come up with challenging questions, here..."

However, two participants were not convinced of the fact that this could make the task of creating tests a more enjoyable one. One of them, for instance, remarked that too many forms had to be filled out in the WoQ-editor, and that this could lead to boredom. For the most positive ones, the majority, all possible increase in enjoyment was subject to the time required and time available; if creating tests with WoQ proved to be a too time consuming task, no amount of enjoyment could compensate for that.

5.8 Behavioral Intention to use the system.

According to the TAM, this is a factor to be deduced from the previous ones. However, we wanted to have a direct measure of this, so we asked participants some questions regarding their intention to use WoQ.

First, they were asked if they would use 3DCVWs in general and WoQ in particular, with assessment purposes in their classes. All of them agreed that the main factor to be taken into account was time: time to learn how to effectively use the system and time to conceive of and implement the questions. The learning curve was generally considered too steep, especially to master the needed operations in the 3D virtual world, and that the conceptual leap required to use the interactive and immersive capabilities of these learning environments should be taken into account. Deep concern was expressed regarding implementation,: how long would it take to obtain the necessary assets (especially 3D objects, but also multimedia content) to implement the test questions they had made?

On the other hand, the perceived need for creativity was considered a motivating factor: all of the participants had the opinion that creative teachers would be able to use WoQ well, no matter what subject they were teaching. This was an insufficient argument, however, for a shared thought that it could be over-elaborate for some technical and mathematical disciplines. One participant, a specialist in accessibility, suggested that the use of WoQ would be very advantageous for elderly learners and special programs for handicapped children and youngsters, as they might find a warmer and more realistic environment in 3DCVWs, where they could feel both better accompanied and better supported.

When asked if they would use WoQ for summative assessment and even for a final exam, they expressed some additional concerns about the security of the computer network. Most of the participants were more inclined to use the environment for formative evaluation.

"I'd rather use it for previous formative evaluation than for giving a mark to my students".

6 Discussion

In general, both the possibility of carrying out tests in 3DCVWs and the use of WoQ to do so were positively welcomed by most participants. Two of them, familiar with the use of 3D environments for educational purposes, stated that they would like a tool for creating and delivering assessment resources in this kind of scenario. Another one enthusiastically suggested going further and more explicitly integrating the gamification potentialities of these 3D virtual worlds into the system. Another two participants, on the other hand, were quite reluctant to change their traditional way of delivering tests in favour of this new approach.

In this section we will take into account all the answers, comments, ideas, feelings and suggestions expressed by the participants in the experiment, in order to answer our research questions, as stated in section 4.

6.1 Do teaching professionals feel confident of their capability to use 3DCVWs in general, and WoQ in particular, for assessment purposes?

According to the participants' answers during the interviews, they definitely do feel confident. They would enthusiastically embrace the task of creating tests for this kind of environment using WoQ as the authoring tool, and that once they had overcome the initial conceptual leap, they would be able to work autonomously. But even though this was the general opinion of the participants, they also expressed some doubts and concerns regarding their performance of the task.

For instance, most of the participants were very critical of the 3DCVW platform user interface, in this case, Open Wonderland. It was by far the most difficult element to use, in their opinion. Operating the interface was awkward and confusing, even if their skills slowly improved as time passed.

Participants also expressed a certain concern about the high level of creativity they perceived was needed. Their feelings tended to be quite contradictory in most instances: they felt excited and considered this need for increased creativity to be a motivating and positive challenge, but at the same time, they felt rather frightened. They wondered if they would be capable of rising to the occasion.

But no doubt the main demand from participants had to do with the explicit management of visibility and groups. Many of them became annoyed when they felt forced to understand the group model implemented in the virtual world in order to be able to include visibility features in their questions when creating them in the Editor. One of the participants wondered if forcing teachers to understand that was right, and that if knowledge of those implementation details was absolutely necessary, then a computer science specialist should collaborate with the them.

6.2 How positively do educators consider the added value in the use of 3DCVWs in general, and WoQ in particular, for assessment? Would they use them for summative evaluation?

In general, most of the participants considered that the value added by these environments and by WoQ to the assessment activities was very important. They felt, for instance, that the features of 3DCVWs and the flexibility and variety of options of

WoQ was tremendously enriching. They especially highlighted the content richness and the interaction capabilities provided by these tools.

However, this perceived enrichment in itself is not enough, and other issues have to be considered, according to the participants. The most important of these is that the value added by these tools is tremendously dependent on the subject being taught and student profiles. 3DCVWs and WoQ have enormous potential, but it could come to nothing if they are used for subject areas that are not well suited for audiovisuals or immersion, or with students who oppose changes in the assessment process. Among the less suitable areas of knowledge, participants cited those related to technology and mathematics, and among the most reluctant learners, the most reading-oriented.

Finally, while most of the participants declared themselves unwilling to make their students' marks depend on this kind of assessment environment, they tended to consider that its use for formative evaluation would be very positive and engaging for their students.

6.3 According to the participants' opinion, what are the main factors influencing their intention of use of 3DCVWs in general and WoQ in particular, for assessment?

As previously stated, two very important factors to take into account when evaluating intention of use of this kind of assessment environment are subject being taught and student profile.

Another factor identified by our study is the learning curve: even if participants feel confident about using the system, they also indicated that the learning curve, including the necessary conceptual leap that must be made to master the system and to take full advantage of its possibilities, could be too steep for some teachers. In their opinion it could make some teachers desist.

But no doubt the most important factor, the one that the participants most frequently mentioned, was time: time to achieve the conceptual leap, time to learn how to operate the system, time to come up with relevant questions, time to get and edit assets and time to implement the tests. Time was raised as an extremely scarce resource for teachers, and even the most enthusiastic stated that they would not be able to use the system if it proved to be too time-consuming, no matter how enriching, relevant, and significant its use could be.

7 Conclusions

In this work we have explored the use of the features of 3D virtual worlds in designing and building assessment systems. These environments are widely used for learning processes, but research on how to include the assessment of learning outcomes in the in-world experience is not equally abundant.

We have tried to take full advantage of the interactive and immersive capabilities of these environments to recreate real or imaginary places to be explored by the students, and have used these 3D settings as working places where assessment takes place in the context of the content under evaluation, i.e., asking about the Football World Championship in a football stadium. This contextualization and immersion are expected to improve student motivation and engagement with assessment activities.

Participants seem confident that their students would be more engaged with the assessment activities when using this kind of environment and therefore they would be able to use the tool for formative assessment.

Despite the limited number of participants in the study, we identified general opinion trends regarding the dimensions of the TAM framework. The evaluations of the WorldOfQuestions authoring system suggests that educators can find it both challenging and frightening to come up with assessment activities that make the most of a 3D environment's possibilities, especially its interactivity. The tool was perceived as useful for including situational assessment activities that would lead to engaging students in learning. However, further studies are necessary to evaluate the impact of these assessment activities on student engagement. Other aspects identified as determinant are the subject being taught and student profiles.

In order to overcome teachers' difficulties in using this and similar assessment tools in 3D virtual worlds, it is advisable to flatten the learning curve by offering teachers appropriate training, and duly supporting them to help them decrease the time needed to develop and deliver tests by using these tools. Furthermore, other implementation approaches should be explored. In this regard, we recommend using platforms that are more stable than Open Wonderland and studying the suitability of integrating the WoQ-editor with a learning management system to make it easier to create groups.

We are currently working to improve the technical reliability of the environment and to automate tasks, such as moving files across the file system and managing visibility, to include new question types, to increase the functionalities and straightforwardness of the WoQ-editor and to raise the collaboration level in the assessing activities. Our next objective is to set up assessment processes of a more deeply constructivist nature.

Our study focused on the deployment of typical closed-choice test questions in 3D virtual worlds, and excluded holistic assessment approaches suitable for measuring complex student knowledge.

Acknowledgements

This research has been partially supported by the Spanish national projects Learn3 (grant TIN2008-05163/TSI) and EEE (grant TIN2011-28308-C03-01) and the Madrid regional project eMadrid (grant S2009/TIC-1650).

References

[Bainbridge 07] Bainbridge, W. S.: "The Scientific Research Potential of Virtual Worlds"; Science, 317, 5837 (2007), 472-476.

[Bell 08] Bell, M.: "Toward a Definition of "Virtual Worlds""; Journal of Virtual World Research, 1, 1 (2008), 1-5.

[Bloomfield 09] Bloomfield, P.R., Livingstone, D.: "Immersive learning and assessment with quizHUD"; Computing and Information Systems Journal, 13 (2009), 20-26.

[Bransford 00] Bransford, J. D., Brown, A. L., Cocking, R. R.: "How People Learn: Brain, Mind, Experience, and School: Expanded Edition"; J. D. Bransford, A. L. Brown, & R. R. Cocking, Eds., Telecommunications Policy 24, (2000), 374.

[Calogne 08] Calongne, C.M.: "Educational Frontiers: Learning in a Virtual World"; Educause Review, 43, 5 (2008), 36-48.

[Chatzopoulou 10] Chatzopoulou, D. I., Economides, A. A.: "Adaptive assessment of student's knowledge in programming courses"; Journal of Computer Assisted Learning, 26, 4 (2010), 258–269.

[Chittaro 07] Chittaro, L., Ranon, R.: "Web3D Technologies in Learning, Education and Training: Motivations, Issues, Opportunities"; Computers & Education, 49, 1 (2007), 3-18.

[Clark 10] Clark, J., Dede, C.: "Assessment, Technology, and Change."; Change 42, 3 (2010), 309-328.

[Dalgarno 04] Dalgarno, B.: "A classification scheme for learner-computer interaction"; in R. Atkinson, C. McBeath, D. Jonas-Dwyer, R. Phillips (Eds), Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference (2004), 240-258.

[Dalgarno 10] Dalgarno, B., Lee, M. J. W.: "What are the learning affordances of 3-D virtual environments?"; British Journal of Educational Technology, 41, 1 (2010).

[Davis 09] Davis, A., Murphy, J., Owens, D., Khazanchi, D., Zigurs, I.: "Avatars, People, and Virtual Worlds: Foundations for Research in Metaverses"; Journal of the Association for Information Systems, 10, 2 (2009), 90-117.

[Delp 07] Delp, S.L., Anderson, F.C., Arnold, A.S., Loan, P., Habib, A., John, C.T., Guendelman, E., Thelen, D.G.: "OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement"; IEEE Transactions on Biomedical Engineering 54 (2007), 1940–1950.

[Dickey 05] Dickey, M.D.: "Brave new (interactive) worlds: A review of the design affordances and constraints of two 3D virtual worlds as interactive learning environments"; Interactive Learning Environments, 13, 1 (2005), 121-137.

[Dillenbourg 02] Dillenbourg, P., Schneider, D., Synteta, P.: "Virtual Learning Environments"; Communication, 8, 6 (2002), 3-18.

[Eschenbrenner 08] Eschenbrenner, B., Nah, F., Siau, K.: "3-D Virtual Worlds in Education: Applications, Benefits, Issues, and Opportunities"; Journal of Database Management, 19 (2008), 91-110.

[Gardner 11] Gardner, M., Gánem-Gutiérrez, A., Scott, J., Horan, B., & Callaghan, V. I. C.: "Immersive education spaces using Open Wonderland: From pedagogy through to practice"; Information Systems Journal, 2 (2011), 1-12.

[Girvan 10] Girvan, C. Savage, T.: "Identifying an appropriate pedagogy for virtual worlds: A Communal Constructivism case study"; Computers & Education, 55, 1 (2010), 342-349.

[Hevner, 2004] Hevner, A.R., March, S.T., Park, J., Ram, S.: "Science in Information Systems Research"; MIS Quarterly, 28, 1 (2004), 75-105.

[Huang 10] Huang, H.M., Rauch, U., Liaw, S.S.: "Investigating learner's attitudes toward virtual reality learning environments: Based on a constructivist approach"; Computers & Education, 55, 3 (2010), 1171-1182.

[Ibáñez 10] Ibáñez, M.B., Crespo, R.M., Delgado Kloos, C.: "Assessment of Knowledge and Competencies in 3D Virtual Worlds: A Proposal"; Proc. KCKS 2010 Key Competencies in the Knowledge Society World Computer Congress Brisbane Australia (2010), 165-176.

[IMS 06] IMS.: IMS question & test interoperability specification v2.0/v2.1. Retrieved on May, 2012, from http://www.imsglobal.org/question/index.html

[Kallonis 10] Kallonis, P., Sampson, D.:" Implementing a 3D Virtual Classroom Simulation for Teachers' Continuing Professional Development"; Proc. of the Workshop on Virtual Worlds for academic, organizational, and life-long learning (ViWo 2010) in the 18th International Conference on Computers in Education (ICCE 2010) (2010).

[Kickmeier-Rust 08] Kickmeier-Rust, M.D., Marte, B., Linek, S., Lalonde, T., Albert, D.: "The Effects of Individualized Feedback in Digital Educational Games"; Proc. of the 2nd European Conference on Games Based Learning, (2008) 227-236.

[Livingstone 08] Livingstone, D., Kemp, J.: "Integrating Web-Based and 3D Learning Environments: Second Life Meets Moodle"; *UPGRADE* The European Journal for the Informatics Professional, IX, 3 (2008), 8-14.

[Mahfujur Rahman 10] Mahfujur Rahman, A. S. Md., Alamgir Hossain, S.K., El-Saddik, A.: "Bridging the Gap between Virtual and Real World by Bringing an Interpersonal Haptic Communication System in Second Life". ISM 2010 (2010) 228-235.

[Mikropoulos 11] Mikropoulos, T. A., Natsis, A.: "Educational virtual environments: A ten-year review of empirical research (1999–2009)"; Computers & Education, 56, 3 (2011), 769-780.

[Minocha 10] Minocha, S., Reeves, A.: "Design of learning spaces in 3D virtual worlds: an empirical investigation of Second Life"; Learning Media and Technology, *35*, 2 (2010), 111-137.

[Nelson 10] Nelson, B.C., Ketelhut, D.J., Schifter, C.: "Exploring Cognitive Load in Immersive Educational Games: The SAVE Science Project", International Journal of Gaming and Computer Mediated Simulations, 2, (2010) 31–39.

[Savin-Baden 08] Savin-Baden, M.: "From cognitive capability to social reform? Shifting perceptions of learning in immersive virtual worlds"; AltJ, 16 ,3 (2008), 151-161.

[Shute 09] Shute, V.J., Ventura, M., Bauer, M., Zapata-Rivera, D.: "Melding the Power of Serious Games and Embedded Assessment to Monitor and Foster Learning: Flow and Grow"; Serious Games Mechanisms and Effects 1, (2009) 1-33.

[Smith 05] Smith, B., Caputi, P.: "Cognitive interference model of computer anxiety: Implications for computer based assessment"; Computers in Human Behavior, 21 (2005), 713–728.

[Terzis 11] Terzis, V., & Economides, A. A.: "Computer based assessment: Gender differences in perceptions and acceptance". Computers in Human Behavior, 27, 6 (2011), 2108-2122.

[Venkatesh 08] Venkatesh, V., Bala, H.: "Technology Acceptance Model 3 and a Research Agenda on Interventions", Decision Sciences 39, 2 (2008), 273–315.