ETeach3D: designing a 3D virtual environment for evaluating the digital competence of pre-service teachers

Abstract: The acquisition of teacher digital competence (TDC) is a key aspect in the initial training of teachers. However, most existing evaluation instruments do not provide sufficient evidence of this teaching competence. In this study we describe the design and development process of a three-dimensional (3D) virtual environment for evaluating the TDC of future teachers, through a performance-based, collaborative and contextual evaluation. This environment, named *ETeach3D*, has been constructed using the Educational Design Research (EDR) approach. It is based on successive iterative cycles and is in accordance with the criteria of usefulness, validity and effectiveness. In addition to the research team responsible for the project, participating in this study were 187 Spanish undergraduate students of Education and 22 experts in the field of educational technology. Results show that these environments, in addition to other characteristics, should: (1) function smoothly and have simple interfaces, realistic scenes and interactive activities; and (2) follow a systematic evaluation procedure that integrates several strategies and levels of complexity. This research helps to improve the initial training of pre-service teachers, and contributes to the growing number of EDR studies that focus in the field of evaluation of the curriculum domain.

Running head: **ETEACH3D: EVALUATING DIGITAL COMPETENCE**

Francesc M. Esteve-Mon, Rovira i Virgili University (*)

Jose María Cela-Ranilla, Rovira i Virgili University

Mercè Gisbert-Cervera, Rovira i Virgili University

(*) Corresponding author information:

Department of Pedagogy, Rovira i Virgili University, Ctra. Valls s/n; 43007 Tarragona, Spain. Phone: +34977558466. e-mail: francescmarc.esteve@urv.cat

Keywords: performance assessment, 3D environments, digital competence, teacher education, educational design research

Dr Francesc M. Esteve-Mon, Dr in Educational Technology, he is lecturer in the Faculty of Educational Sciences and Psychology at Rovira i Virgili University (URV). He teaches in Initial Teacher Education, Pedagogy and Educational Technology programs. He participates in the Applied Research Group in Educational Technology (ARGET), and develops his research on digital competences, teacher training and new virtual environments.

Dr Jose María Cela-Ranilla is a lecturer at the Faculty of Educational Sciences and Psychology of the Rovira i Virgili University in Tarragona, Spain. He is a member of the Applied Research Group in Educational Technology (ARGET), which conducts research on the use of ICT in Education, where he has participated in several national and international projects. He is particularly interested in the application of Educational Research Methodology (EDR).

Dr Mercè Gisbert-Cervera has a permanent position as professor at Rovira i Virgili University in the Department of Pedagogy. She is Dr in Educational Sciences and since 1988 has been conducting academic and research activities. She is the coordinator of the interdisciplinary Applied Research Group in Technology and Education (ARGET), and responsible of a PhD and Master degree on Technology and Education.

1. Introduction

The role of teachers is crucial to help and empower students with the advantages of the technology in today's society (UNESCO, 2008). Teachers need to be prepared to use the Information and Communication Technology (ICT), and provide technology-supported learning opportunities for their students (European Commission, 2013). This digital competence (DC) needed by teachers has two dimensions: (1) mobilise knowledge, abilities and attitudes in order to use ICT efficiently; and (2) improve and transform classroom practices and enrich the professional development and identity of both teachers and students (Hall, Atkins & Fraser, 2014; Krumsvik, 2009). Initial teacher education should enable student to achieve the teacher digital competence (TDC) (Gutiérrez, Palacios & Torrego, 2010) and to this purpose, universities have to reflect on what are the most suitable strategies for teaching and evaluating DC of pre-service teachers (Redecker, 2013).

The technological advances of the last few decades are enabling innovative forms of evaluation based on assessing the student's performance on a series of learning experiences (Clarke & Dede, 2010; Code, Clarke-Midura, Zap & Dede, 2013). Specifically, 3D technology provides technological advances that enable new forms of active and contextualized learning and evaluation to be designed and developed, as well as in-depth observations of the student's learning process to be made (Andrews & Wulfeck, 2014). These types of practices are now beginning to be used in teacher training (Chau, Wong, Wang, Lai, Chan, Li, Chu, Chan & Sung, 2013; Christensen, Knezek, Tyler-Wood & Gubson, 2011; Gregory, Dalgarno, Crisp, Reiners, Masters, Dreher & Knox, 2013).

This type of evaluation, based on performance and measured by ICT, opens up a wide range of possibilities for teacher development. However, several authors have shown that, so far, not all of its potential has been accessed (Clarke & Dede, 2010), especially in the Social Sciences (Kuo & Wu, 2013). In the case of TDC, for example, most evaluation instruments are paper-based or use simple computer software that does not cover the whole complexity of this competence or the student's performance of it (Esteve-Mon & Gisbert-Cervera, 2013). Moreover, many of these tools have not been designed as a part of teaching-learning processes and providing proper feedback and possibilities to further improvement, that is to say from a formative perspective. Contrarily, most of them have been designed by extern institutions oriented to certifying processes (Esteve-Mon, 2015).

In this study, we developed a 3D virtual environment (ETeach3D) for evaluating the DC of pre-service teachers, through a performance-based, collaborative and contextual evaluation, and as part of an instructional process. An educational design research (EDR) approach was utilized over the three years of this study (van den Akker, Gravemeijer, McKenney & Nieveen, 2006). The purpose of this article was (1) to explain the iterative process of design, development and evaluation that was used to create this environment, and (2) to describe design principles formulated during the process.

2. The digital competence of pre-service teachers

The term "digital competence" has evolved over the last few decades, though it has always been associated with the various literacies of the new media (Lankshear & Knobel, 2008). Digital literacy comprises aspects such as the identification and treatment of information, the creation of content, communication, and the safe use of digital tools (Covello, 2010; Gilster, 1997).

Although "digital literacy" is the term most widely used internationally, often, and especially in European contexts, the term "digital competence" (Ferrari, 2012) is used synonymously. According to Ferrari (2012), digital competence implies not only having certain abilities, knowledge and attitudes, but also the capacity to put these in action and mobilise them in a certain context.

Primary and secondary school teacher trainers require not only basic digital literacy; they also need to be able to incorporate technology into their teaching praxis (Krumsvik, 2008). As Hall, Atkins and Fraser, (2014) have suggested, a digitally competent teacher is one who possesses the abilities, attitudes and knowledge that are needed to engender true learning in a context that is enhanced by technology. Teachers must therefore be able to use technology in order to improve and transform their classroom practices and to enhance both their own identity and professional development as well as those of their students (Redecker, Ala-Mutka, Bacigalupo, Ferrari & Punie, 2009).

Krumsvik (2009) defines this competence on several levels and according to several key competences: (a) basic ICT competence, i.e. the knowledge and skills required to access information and communicate in everyday situations; (b) ICT teaching competence, i.e. the ability to use digital tools together with suitable teaching strategies to enable the acquisition and construction of knowledge; and (c) learning strategies, i.e. the resources and tools that enable the user to learn continuously. Similar to this model is the one proposed by Kabakçi (2009), which additionally includes aspects related to knowledge transfer and management.

To guide the process of training and evaluation of the digital competence of current and future teachers, various administrations and institutions have developed their own frameworks for performance standards and indicators. Two types of model exist: the first one focuses more on basic digital skills while the second is more holistic and focuses on integrating ICTs into teaching and learning processes (Silva, 2012). The International Computer Driving License (ICDL) is an example of the first type of model, which focuses on the basic use of ICT and includes aspects such file management, word processing, spreadsheets, and presentations tools. Other models, such as the UNESCO ICT Competency Framework for Teachers (UNESCO, 2008), the National Educational Technology Standards for Teachers (NETS-T) of the ISTE (International Society for Technology in Education) (ISTE, 2008) and Enlaces (2011), focus on the application of ICT in training processes. They include aspects that are more related to the teaching and learning process (e.g. the design of learning experiences and evaluations that include ICTs), the teacher's professional development via ICTs, as well as institutional management and the socio-educational context.

We used the NETS-T (ISTE, 2008) standards as reference for the design and development of the 3D virtual environment of this research. These standards represent a holistic and cross-disciplinary model that approaches a constructivist vision of education (Morphew, 2012). It separates the indicators into four levels of performance (beginner, intermediate, expert, and transformer) and divides them into five dimensions: (student learning and creativity through the use of ICTs; (2) student learning experiences and assessments via technology; (3) digital-age work and learning; (4) digital citizenship and responsibility; and (5) professional growth and leadership through digital tools.

3. Evaluating digital competence using 3D virtual environments

Despite the diversity of TDC frameworks and models, according to Esteve-Mon & Gisbert-Cervera (2013) most evaluation tools do not adequately cover every aspect of this competence. Firstly, many of these tools only focus on analysing basic digital skills, i.e. the appropriate and efficient use of various software and hardware rather than their application to the teaching profession. And secondly, many existing instruments are either paper-based or simple computer simulations that do not lend themselves to the performance or evaluation of complex activities. Furthermore, most of these have been designed from a certifying and external perspective, as objective testing programs. These programs try to reach the improvement and accountability by means of grading and ranking; these do not consider a formative and constructivist view integrated in learning and teaching processes (Esteve-Mon, 2015; Stufflebeam, 2002).

In the last ten years, several advanced technological environments have appeared that are especially suited to the development and evaluation of competences (Redecker, 2013). One example of these are 3D virtual environments, also called, though with one or two slight differences in meaning, metaverses or multi-user virtual environments (MUVEs) (De Freitas, 2008). 3D virtual environments, such as Second Life and OpenSim, are online communities that simulate physical spaces in three dimensions that may or may not be similar to real spaces. Via avatars, they allow users to interact with each other and with the environment, and to use, create and exchange objects.

These environments can be immersive, interactive, personalisable, accessible and programmable (Atkins, 2009), and have numerous potential uses for educational praxis and research (Cela-Ranilla, Esteve-González, Esteve-Mon & Gisbert-Cervera, 2014; Dalgarno & Lee, 2010). For the interaction and immersion sensation to be effective and attractive, an intuitive navigation system is required with a sequence of activities and clear instructions that the user can understand and follow (Eseryel, Guo & Law, 2012). These aspects, plus the ability to communicate and collaborate easily with other users, realism and the quality of the sensorial (visual, auditory and tactile) stimulants, are highly motivating for the user (Olasoji & Henderson-Begg, 2010; Wilson, Bedwell, Lazzara, Salas, Burke, Estock, Orvis & Conkey, 2009).

In the last few years, the characteristics and potential of these environments have led to the development of evaluation experiences that use immersive and 3D-simulation technology and are based on student performance (Code et al., 2013). As well as creating a suitable ambiance for didactic activities, 3D environments enable the students' actions and behaviours during these activities to be collected automatically and non-intrusively, thus enabling multiple tests and methods to be integrated in a single evaluation in a practical, valid and viable way (Clarke & Dede, 2010; Clarke, Code, Zap & Dede, 2011).

To ensure that these evaluations are suitably developed, simulated environments must be designed that combine every feature of these environments in a valid way, including features pertaining to competence, the didactic sequence/problem to solve and aspects of evaluation as well as those pertaining to the software and the design of the environment (Mislevy, 2011). To achieve this, evaluation planning models such as evidence-centred design (ECD) enable these processes to be systemised and provide valid evidence of the learning achieved or competence acquired. This model, which has been implemented in several evaluations of 3D learning environments (Nelson, Ketelhut, Clarke, Bowman & Dede, 2005; Shute, Masduki & Donmez, 2010), will be used in the present study.

Despite their potential for creating evaluation experiences, these environments also present certain limitations. Some authors have stressed that the use of these technologies in evaluation is not without certain problems (Olasoji & Henderson-Begg, 2010). For example, users may find them difficult to use because of a high learning curve associated with the tool, while creating these evaluation scenarios may be a laborious process. Moreover, despite the sensorial and situational complexity of this 3D technology, the number of possible actions or interactions that can be recorded and from which researchers can collect data, is limited (Nelson, Erlandson & Denham, 2011).

In the last ten years, several experiments in 3D environments have been developed at the pre-university level (Ketelhut, 2007; Nelson et al., 2005; Quellmalz, Silberglitt & Timms, 2011). Some of these, including River City and SimScientist, have followed the EDR-DBR method and the ECD model. Others have been developed for the initial training of primary school teachers. Woollard and Wankel (2011) conducted an experiment using Second Life for students of Education, with positive results regarding the acquisition of teaching-related cognitive, procedural and social aspects. The results of other studies using 3D virtual environments highlight the usefulness of these technologies to develop several competences of future teachers (Christensen et al., 2011; Gregory et al., 2013; Sparrow, Blevins & Brenner, 2011). However, despite the close relationship between those experiments and the subject of the investigation, none of the existing environments completely matched the purpose or context of the present study. For this reason, and using the above experiments as a starting point, we decided to develop a new environment with which to evaluate TDC.

Based on research works by McKenney & Reeves (2012) and Dowse & Howie (2013), the basic concepts already mentioned in the theoretical section were the starting point to define a logic model for the present work. Table 1 shows visually the relationship between the inputs, the planned processes, and the expected results.

-Insert Table 1 here-

4. Research design

This study used an Educational Design Research (EDR) approach, which is a variant of Design-Based Research (DBR) applied to Education (van den Akker et al., 2006). According to Plomp and Nieveen (2009), EDR studies involve a systematic process for designing, developing and evaluating an educational intervention as a solution to a complex problem that is often technology-related.

These complex educational processes are usually related to some curricular components. According to McKenney, Nieveen and van den Akker (2006) and in accordance with van den Akker (2003), the assessment systems are one of the ten interconnected components of the curriculum that also include educational materials, contents and learning activities. This design-based research focuses on the development of a 3D environment for evaluating digital competence within a formal educative process, and is therefore another essential component of the teacher-training system. Moreover, by its nature this type of EDR study: (1) is characterised by interaction and collaboration among the various stakeholders involved in the training process (researchers, teachers and students); and (2) although it focuses on designing and developing an educative intervention for a specific context, follows quality criteria and a systematic process for analysing these phenomena (Plomp and Nieveen, 2009) that we will now describe in greater detail.

The aims of the intervention are not only to find a solution to the problem in hand and increase the knowledge but also to generate a series of design principles that can be applied to other situations. In this study we have conducted a process that is structured in three phases: (1) preliminary, (2) iterative design, and (3) assessment. Figure 1 presents and illustrates the research process as a whole. However, all the phases of the study are explained in detail in this paper.

-Insert Figure 1 here-

In the preliminary phase we reviewed and analysed the literature. This helped to establish the conceptual basis for the study with regard to TDC (Richey & Klein, 2005). In this phase, we also analysed the context further by evaluating the perception university students have of their own digital competence.

In the second phase we conducted an iterative process for designing, developing, evaluating and reviewing several prototypes (Collins, Joseph & Bielaczyc, 2004; McKenney, 2001) of a 3D environment for evaluating TDC. This environment had to be a complete system that would enable TDC to be deployed in such a way that evidence of the student's performance could be collected in a valid and systematic way and in line with the definition outlined in the theoretical framework. Different internal and external experts analysed the iterative process in accordance with quality criteria for the technological and graphical usability of the environment, the validity of content and appearance, pedagogical practicality and usefulness, and the effectiveness of the system.

The final phase of the process has an important final summative evaluation component, which is intended to analyse the efficiency of the process as a whole (Plomp & Nieveen, 2009). Also in this phase several documents were drawn up and, in line with EDR approach, the design principles extracted from all iterations were produced in order to help researchers prepare future proposals with similar situations but in different contexts, and they are described in the results section of this article.

4.1 The 3D virtual environment (ETeach3D)

Following the approach described above, we designed a complex 3D environment to assess the pre-service teachers' digital competence. From the technological perspective, the 3D environment was created using the free OpenSim software package, which enables virtual worlds to be created and configured, and it was linked with Moodle through a Sloodle plugin to identify the students and record all activities. To enable student access to the environment, a virtual world viewer was installed and configured in two computer laboratories at the School of Education. From the graphical point of view, several ad-hoc scenes were designed, and textures and objects from open repositories were incorporated. This virtual environment was intended to simulate a primary school designed from authentic examples from the real context under investigation.

From the conceptual perspective, we took as references the ECD model and the ISTE NETS-T international standards of digital competence for teachers. This framework served as the conceptual basis for the design of the scenes (SCn) and activities (An) and for the evaluation procedures. The three main scenes were: (SC1) "The classroom, spaces and resources", which simulated a primary school classroom with its furniture and other resources; (SC2) "The didactic activity workshop", which simulated a multi-purpose room in which to design didactic activities; and (SC3) "The staff room", which simulated the teachers' work space and meeting room and had additional area for continuous training activities and meetings with families. There are six activities, two in each scene. These activities can be summarized as follows: (A1) discuss and reorganize collaboratively the physical learning environments and classroom materials; (A2) select and justify an array of complementary ICT resources; (A3) discuss ways of using a technology found in the scene to locate, analyse and create certain digital products; (A4) design the learning activities proposed above and find websites with resources; (A5) reply to certain messages received on the simulated school computers; and (A6) discuss how to use the technology of the simulated staff room for working with colleagues.

5. Method

ETeach3D construction process was iterative and, as suggested by other authors (Dede, Nelson, Ketelhut, Clarke & Bowman, 2004; Tessmer, 1993), its cycle combines quantitative and qualitative methods for obtaining important information systematically with the participation of key informants and potential users.

5.1 Empirical context

The study, conducted between 2012 and 2014, comprised four iterations for the refinement of the ETeach3D intervention, with different instruments, participants and data-collection procedures. Participating in the study were 187 third- and fourth-year Spanish university students of Primary Education Teacher Training and Pedagogy and 22 experts from the field of Educational Technology, six of whom had a technology profile and 16 had an academic profile (university lecturers and pre-university teachers). The participants were divided among the three phases of the process (see Figure 1).

5.2 Instruments

To collect the information generated in the various iterative cycles, the researchers used the following instruments:

TDCSQ (Teacher Digital Competence Self-Perception Questionnaire)

This is a questionnaire of 40 items with a continuous Likert scale of 1 to 8, where 1 = not proficient at all and 8 = highly proficient. The questionnaire was constructed in accordance with ISTE standards, validated by a sample of experts, and tested for reliability (alpha = 0.96) in the sample (Esteve-Mon, 2015). It was applied to a sample of 149 university students from the field of Education.

Control list

To detect obvious errors in the initial paper prototype design, the researchers used a dichotomous-response control list comprising the following elements formulated in accordance with Tessmer (1993):

a) Content quality: Content accuracy, content currency, content completeness and content superfluousness.

b) Learner performance: Clarity of writing, proper sequencing of content, effectiveness of strategies, realistic examples, workplace performance and quality of feedback.

c) Learner interest: Interest in content, level of learner challenge, perceived value of learning and time spent learning.

d) Implementability: Teacher ease to use, learner ease to use, orientation requirements, and support requirements.

Discussion groups

The second prototype was analysed in 90-minute sessions with two discussion groups. The first group comprised experts in technology (n = 6) and the second comprised undergraduate students of Education (n = 10). Both groups analysed the technological and graphical usability of the environment. At these sessions the participants examined the 3D environment and, guided by the researchers, provided their collective opinions on the following topics: (a) the technical quality of the environment (i.e. image and sound quality, the performance of the hardware and software, possible problems with the technology, and the suitability of the tools, etc.); and (b) user control (i.e. avatar movements and interactions with the scene, objects, and other users, etc.).

Content validity questionnaire (CV)

The CV questionnaire (Esteve-Mon, Adell-Segura & Gisbert-Cervera, 2014) contained five items for validating the following elements: content adaptation, realism, topical interest, clarity, and time allowed. The items were evaluated on a Likert scale, where 1 = completely disagree and 5 = completely agree.

Pedagogical Usefulness questionnaire (PU)

This questionnaire was adapted from Code et al. (2013) and in line with Nokelainen (2006). It has a Likert scale, where 1 = completely disagree and 5 = completely agree, for the following components of the 3D environment: code comprehensibility, user control, reflexive thinking, immersion sensation, communication and dialogue, teamwork, perception of usability, added value of content, added value of graphical interface, interest for studies, extrinsic motivation, and valuation of previous knowledge. The questionnaire was completed by 28 students (as potential end users) after surfing and interacting with the scenes and activities of the ETeach3D for 120 minutes.

Evaluation rubric (NETS-T of the ISTE)

The fourth ETeach3D prototype was administered to a sample of 13 university students of Primary Education and Pedagogy. Taking the NETS-T rubric of the ISTE as reference, a group of experts applied a control list to evaluate the performance of the students during the session in the assessment activities described above and with the elements recorded on the server. This control list comprised a 4-point Likert scale, where 1 = poor, 2 = fair, 3 = good and 4 = excellent.

Table 2 summarizes the data collection instrument used in each phase, the criteria and the participants.

-Insert Table 2 here-

6. Results

From the perspective of EDR approach, each phase of the process provides important and interesting results that are worth sharing. Therefore, to deliver an ordered and comprehensible account, we will report the results from each phase of the ETeach3D construction process outlined in the study design.

6.1 Preliminary phase

Once we had analysed the context and reviewed the literature, we took the ISTE indicators (2008) as the conceptual framework of reference for the systematisation of the evaluation process. These indicators were used to develop the process that began with the first analysis of the students' self-perception of TDC.

Table 3 shows that the vast majority of students considered themselves quite capable, or very capable, of displaying teacher digital competence (average score = 6.11).

-Insert Table 3 here-

Table 3 also shows that the students perceived that they have more competence in dimensions related to the use of technology (dimensions 3 and 4, with average scores of 6.17 and 6.49, respectively). They perceived that they have less competence in dimensions related to technology applied to teaching (dimensions 1 and 2, with average scores of 6.04 and 6.11, respectively) and especially in the use of ICT for continuous professional growth (dimension 5, with an average score of 5.74). These results are in line with those reported by Banister and Reinhart (2012) and Oh and French (2004).

6.2 Prototyping phase

Iteration 1: Paper-based prototype

The first prototype was a theoretical (i.e. paper) model that integrated all the gathered elements from the review of the literature and the analysis of the context. We then used these elements to create a hypothetical virtual environment comprising several scenes.

Iteration 2: Computer-based prototype

The second prototype comprised the first of the three 3D-simulation environments (see Figure 2) and incorporated the contents of the evaluation activities. The scenes were designed in rudimentary fashion, some basic objects were created, and the instructions for the activities were included.

-Insert Figure 2 here-

The results of the technological analysis obtained from the focus group showed that, despite the highly demanding requirement for equipment performance, the environment was stable and fluid and the information presented was practical and easy to understand. However, it was also reported that the graphics for this prototype were too primitive and basic, especially compared with those of computer videogames. Another aspect that was less well-evaluated by both experts and students was user control. These aspects were corrected for later versions.

The content analysis conducted by the experts using the CV questionnaire provided values that validated content adaptation, topical interest and the realism of the activities. The aspects that were evaluated the lowest were related to the clarity of the instructions given to the students and the time they were given to complete them. We created the next 3D learning environment considering the results and all comments received.

Iteration 3: Entire system prototype

After the above improvements had been incorporated (see Figure 3), the results showed that this 3D environment prototype was especially motivating for students both regarding the technology used and the content of the activities they completed. According to the students, these activities were very useful for practising the real abilities they will need in their future professional careers.

-Insert Figure 3 here-

Generally speaking, all items received a positive average score. Those that received the highest scores from the students were intrinsic and extrinsic motivation (4.00 and 4.11 out of 5, respectively). Those that received the lowest scores were user control (3.11) and added value of graphics (3.21). Both of these areas were improved for the next prototype (see Figure 4).

-Insert Figure 4 here-

The results of the *PU* questionnaire also show that high scores were awarded for immersion sensation (students were so immersed in the environment that they forgot both the time and the activities going on around them), communication with colleagues, and teamwork. Also highlighted were the value of the activities in terms of interest in the topic, the fostering of reflexive thinking, and the motivation and competitiveness for completing the activities correctly. As with the previous iterations, the results of this analysis served as the basis and rationale for the modifications and improvements applied to the next prototype.

6.3 Assessment phase

Iteration 4 – Final version of ETeach3D

Finally, after incorporating the improvements derived from the cyclical process of the iterations, we designed a new version of the ETeach3D environment. In this final version, the environment was applied in a real situation in order to evaluate the digital competence of a small sample of undergraduate Education students (Figure 5).

-Insert Figure 5 here-

The results indicate that the majority of students in teacher training (7 of the 13 participants) achieved a moderate level of digital competence. The best results were for basic digital competences. However, the results were noticeably lower for competence in the didactic use of ICT and in strategies to enable the continued use of digital tools in their professional development. To compare the validity of the results and, therefore, the effectiveness of this evaluation, we used the results of the students' self-perception questionnaire (see Figure 6).

-Insert Figure 6 here-

The self-perception scores were higher in all dimensions. However, the fact that both sets of results follow a similar trend entail a first estimate for this criterion validity.

6.4 Design principles

Van den Akker (2002) suggests that the knowledge obtained from a design and development study such as this one can be conveyed through a series of design principles, or heuristic statements, that serve as a guide for future studies. Adapting these guidelines, in line with similar studies (McKenney, 2001; Zulkardi, 2002), and taking the outputs from the logic model used in this study (see Table 1), we recommend anyone wishing to create a 3D environment for evaluating the teaching competences (especially digital competences) of university students of Education to apply the principles:

- *ETeach3D validation*: (1) Have the bandwidth capacity and speed, and suitable computers to ensure the system functions correctly; (2) use simple 3D viewer with "clean" and intuitive graphical interfaces that are not require a high learning curve; (3) have a scene available for testing the environment in order to enable the users to familiarise themselves with the interface and the environment and to interact with the objects and other avatars, (4) design realistic scenes that take into account not only the main objects of the action but also secondary details of "decoration" and incorporate new textures and sophisticated objects in order to improve the user's immersion sensation, and (5) incorporate components of gamification, which promote a sense of competition between the users and raise the students' extrinsic motivation to use the 3D environment. All these principles were obtained from the first and second iterations of the intervention.

- *Formative evaluation*: (1) design activities that enable interaction and communication between users. This promotes joint reflection and helps the students to become more immersed in the environment; (2) establish an evaluation system that enables observation of the knowledge or competence tested to become operative via evidence, tasks or situations and its interpretation to become systematised; (3) use valid, internationally recognised models to define the competence to be evaluated. This makes the content more relevant and valid and enables the results to be compared with those from similar studies; (4) simulate environments and activities that are similar to real professional ones the students are likely to meet, taking as reference contexts and authentic activities that are familiar to the user; (5) combine several types of evaluation activities and strategies with various levels of complexity depending on the cognitive load and using different instruments to record information on knowledge, comprehension, behaviour and performance. All these principles were obtained from the third and fourth iterations of the intervention.

7. Discussion

This study has been produced within the framework of the educational design research to create a 3D virtual environment for evaluating the digital competence of pre-service teachers. It derives from the need to have a suitable contextualized assessing system in light of the growing need for the teachers of the future to develop their digital competence. The research consisted of an iterative process for designing, developing and revising ETeach3D (see Figure 1), and this process generated the two main outputs that were initially defined in Table 1 and which now serve as the structural basis for the Discussion section of this article.

Regarding the first output, results indicate that graphically sophisticated scenes with at least a minimum level of realism need to be designed in order to ensure an adequate immersion sensation for the user. Results also show that the activities designed need to allow for interaction, communication and competition since they improve the user's immersion sensation and motivation, which, as Eseryel et al. (2012) and Wilson et al. (2009) have indicated, are important and powerful educational characteristics for such environments. However, these graphics must not compromise system performance too much or make the interface too complex or the user control too difficult. These results are in line with those of Quellmalz, Timms, Silberglitt and Buckley (2012), who highlight the disjunction that exists between the graphical quality and the high level of graphical power needed to allow the system to run smoothly. As suggested by Olasoji and Henderson-Begg (2010), despite the proliferation of videogames, the learning curve for this type of 3D technology should be borne in mind and initial learning mechanisms for these tools should be established.

Moreover, results indicate that the instructions must be clear and direct in order to facilitate comprehension and that the activities should be similar to authentic ones. This helps to raise users' perceptions about the usefulness of the environment because it enables them to practise skills they will need in their professional careers. These results are in line with those of Clarke and Dede, (2009), Dalgarno and Lee (2010), and Gregory et al. (2013).

Regarding the second output, on the one hand results indicate that, as Code et al. (2011) and Shute, Masduki and Donmez (2010) have pointed out a systematic evaluation procedure such as the ECD model enables an evaluation to be conducted that is valid, effective and rigorous. As Roelofs and Sanders (2007) have also suggested, various levels of complexity should be established for the proposed activities and, as Rodríguez Espinar and Prades (2009) and Clarke and Dede (2010) have also indicated, several types of evaluation activities and strategies should be used in order to generate more precise observations of student performance in this competence. On the other hand, these students achieved a moderate level of TDC, with better scores in the use of the tools and lower ones for aspects related to their didactic use. These results are in line with those of Almås and Krumsvik (2007) and Gutiérrez et al. (2010). However, these results do present some limitations due to the small sample size for the final iteration. Although we can use these results to analyse local context, they do not enable these results to be generalised.

8. Conclusion

This study enabled us to visualise the process followed for the creation of a 3D environment for evaluating TDC. Despite their limitations and the caution required due to the characteristics of the study, the results provide important information to enable Schools of Education to continue improving the training they provide and, therefore, to continue improving the teaching and learning of future generations, which is undoubtedly one of the most important challenges of the 21st century. These results also enable us to make a series of recommendations for other studies that have similar objectives.

If competence is displayed by actions, it is by actions that it must be evaluated. The design of an environment in which students have to act, try, fail, redo, etc. places this study on the path towards conceptual coherence. Indeed, the iterative cycles presented here always contemplate the action component.

As it was mentioned above, this study involves the first iterative cycles in terms of

EDR. These first steps constitute a design phase of an environment in which the main objectives to be aimed are the internal robustness and getting a local impact. Therefore, the instruments and strategies used for data collection as well as the further analysis were oriented to describe the context, and to validate the consistency and practicality criteria.

We are aware that not having evidences about the possible transfer of the designed environment could be considered a kind of limitation. However, getting evidences about the local and internal criteria is the base to analyse a broad impact in terms of EDR. If we consider that limitations are intrinsic to any research study, we may say that being able to identify them is a clear indicator that the study has been conducted in a rigorous manner. The value added to the observed limitation is precisely to transform it into a proposal for another study and in this way visualise its continuity. In this context, it would be interesting to generate research processes in which the transferability of the resulting design principles were verified. This possible continuation would enable our proposal, which ends with just a local impact, to evolve, generating possible implications and/or adaptations in different contexts that would provide information about its global nature.

From this point forward, we will consider new lines of work for the future studies. On the one hand, we will investigate using ETeach3D with larger samples and in a variety of contexts in order to verify its transferability, or at least its adaptability. At this point, it will be appropriate to propose instruments and analysis focused on psychometric properties to measure the broader impact, making inferences with a larger sample size beyond the descriptive analysis.

And on the other hand, we will use this simulated school not only to evaluate TDC but also to teach it and acquire it by creating a series of activities based on real experiences that can be simulated in this 3D environment. Finally, we will use this environment to develop not only teacher digital competence but other teacher competences too.

In summary, this paper contributes to the growing number of educational design research studies that focus in the curriculum domain. According to the classification of McKenney et al. (2006), many of these studies concentrate on designing and developing new educational materials, learning activities or contents. In this study, however, we have explored new possibilities for research in the field of formative evaluation. As well as being novel, this field is also both enormously complex and important for educational research.

Acknowledgements

This research is funded by the Ministry of Economy and Competitiveness of Spain, SIMUL@B (ref: EDU2013-42223-P), and supported by the Secretary of Universities and Research, Department of Economy and Knowledge, Government of Catalonia. The project is coordinated by the ARGET research group (ref: 2014SGR1399).

References

Almås, A. G., & Krumsvik, R. (2007). Digitally literate teachers in leading edge schools in norway. *Journal of In-Service Education*, *33*(4), 479-497. doi:10.1080/13674580701687864

Andrews, D.H., & Wulfeck, W.H. (2014). Performance Assessment: Something old, something new. In J.M. Spector, M.D. Merrill, J. Elen, & M.J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 303-310). New York: Springer.

Atkins, C. (2009). Virtual experience: Observations on second life. In M. Purvis & B. Savarimuthu (Eds.), *Lecture Notes in Computer Science: Computer-Mediated social networking* (pp. 7-17). Berlin: Springer. doi:10.1007/978-3-642-02276-0_2

Banister, S., & Reinhart, R. (2012). Assessing NETS-T performance in teacher candidates: Exploring the wayfind teacher assessment. *Journal of Digital Learning in Teacher Education*, *29*(2), 59-65.

Chau, M., Wong, A., Wang, M., Lai, S., Chan, K. W., Li, T. M., Chu, D., Chan, I. K. W., & Sung, W. -K. (2013). Using 3D virtual environments to facilitate students in constructivist learning. *Decision Support Systems, 56*, 115-121. doi:10.1016/j.dss.2013.05.009

Christensen, R., Knezek, G., Tyler-Wood, T., & Gubson, D. (2011). SimSchool: An online dynamic simulator for enhancing teacher preparation. *International Journal of Learning Technology, 6*(2), 201-220.

Cela-Ranilla, J. M., Esteve-Gonzalez, V., Esteve-Mon, F., & Gisbert-Cervera, M. (2014). 3D simulation as a learning environment for acquiring the skill of self-management: An experience involving spanish university students of education. *Journal of Educational Computing Research*, *51*(3), 295-309. doi:10.2190/ec.51.3.b

Clarke, J., & Dede, C. (2010). Assessment, technology, and change. *Journal of Research on Technology in Education, 42*(3), 309-328.

Clarke, J., Code, J., Zap, N., & Dede, C. (2011). Exploring the utility of a virtual performance assessment. In *SREE fall 2011, Advancing Education Research*.

Code, J., Clarke-Midura, J., Zap, N., & Dede, C. (2011). Virtual performance assessment in immersive virtual environments. In *Interactivity in e-learning: Case studies and frameworks* (pp. 230-252). IGI Global. doi:10.4018/978-1-61350-441-3.ch011

Code, J., Clarke-Midura, J., Zap, N., & Dede, C. (2013). The utility of using immersive virtual environments for the assessment of science inquiry learning. *Journal of Interactive Learning Research, 24*(4), 371- 396.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *Journal of the Learning Sciences, 13*(1), 15-42. doi:10.1207/s15327809jls1301_2

Covello, S. (2010). *A review of digital literacy assessment instruments.* New York: Syracuse University, School of Education.

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3- D virtual environments? *British Journal of Educational Technology, 41*(1), 10-32.

De Freitas, S. (2008). *Serious virtual worlds. A scoping guide.* UK: JISC e- Learning Programme, The Joint Information Systems Committee (JISC).

Dede, C., Nelson, B., Ketelhut, D. J., Clarke, J., & Bowman, C. (2004). Design- based research strategies for studying situated learning in a multi- user virtual environment. In *Proceedings of the 6th international conference on learning sciences* (pp. 158-165).

Enlaces. (2011). *Competencias y estándares TIC para la profesión docente. Centro de Educación y Tecnología (Enlaces).* Ministerio de Educación, Gobierno de Chile.

Eseryel, D., Guo, Y., & Law, V. (2012). Interactivity design and assessment framework for educational games to promote motivation and complex problem-solving skills. In *Assessment in game-based learning* (pp. 257-285). New York: Springer. doi:10.1007/978-1-4614-3546- 4_14

Esteve-Mon, F.M., & Gisbert-Cervera, M. (2013). La competencia digital en la educación superior: Instrumentos de evaluación y nuevos entornos. *Enl@ce: Revista Venezolana De Información, Tecnología y Conocimiento*, *10*(3), 29-43.

Esteve-Mon, F. M., Adell-Segura, J., & Gisbert-Cervera, M. (2014). Diseño de un entorno 3D para el desarrollo de la competencia digital docente en estudiantes universitarios: Usabilidad, adecuación y percepción de utilidad. *RELATEC, Revista Latinoamericana de Tecnología Educativa*, *13*(2), 35-47.

Esteve-Mon, F.M. (2015). *La competencia digital docente: Análisis de la* *autopercepción y evaluación del desempeño de los estudiantes universitarios de educación por medio de un entorno 3D* (Doctoral dissertation). Rovira i Virgili University, Tarragona.

European Commission. (2013). *Analysis and mapping of innovative teaching and learning for all through new Technologies and Open Educational Resources in Europe. SWD(2013) 341 final.* Brussels.

Ferrari, A. (2012). *Digital competence in practice: An analysis of frameworks.* Sevilla: European Commission, Joint Research Centre (JRC).

Gilster, P. (1997). *Digital literacy.* New York: Wiley Computer.

Gregory, S., Dalgarno, B., Crisp, G., Reiners, T., Masters, Y., Dreher, H., & Knox, V. (2013). *VirtualPREX: Innovative assessment using a 3D virtual world with pre-service teachers.* Sydney: Office for Learning and Teaching, Australian Government.

Gutiérrez, A., Palacios, A., & Torrego, L. (2010). La formación de los futuros maestros y la integración de las TIC en la educación: Anatomía de un desencuentro. *Revista de Educación*, *352.*

Hall, R., Atkins, L., & Fraser, J. (2014). Defining a self-evaluation digital literacy framework for secondary educators: The digilit lecister project. *Research in Learning Technology*, 22. doi:http://dx.doi.org/10.3402/rlt.v22.21440

ISTE. (2008). *National educational technology standards for teachers. Washington* DC: International Society for Technology in Education.

Kabakçi, I. (2009). A proposal of framework for professional development of turkish teachers with respect to information and communication technologies. *Turkish Online Journal of Distance Education, 10(*3).

Ketelhut, D. J. (2007). The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in river city, a multi-user virtual environment. *Journal of Science Education and Technology, 16*(1), 99- 111. doi:10.1007/s10956-006-9038-y

Krumsvik, R. (2008). Situated learning and teachers' digital competence. *Education and Information Technologies, 13*(13), 279-290. doi:10.1007/s10639-008-9069-5

Krumsvik, R. (2009). Situated learning in the network society and the digitised school. *European Journal of Teacher Education, 32*(2), 167-185. doi:10.1080/02619760802457224

Kuo, C. -Y., & Wu, H. -K. (2013). Toward an integrated model for designing assessment systems: An analysis of the current status of computer- based assessments in science. *Computers & Education, 68*, 388-403. doi:10.1016/j.compedu.2013.06.002

Lankshear, C., & Knobel, M. (2008). *Digital literacies: Concepts, policies and practices.* New York: Peter Lang.

McKenney, S. (2001). *Computer-based support for science education materials developers in Africa: Exploring potentials* (Doctoral dissertation). University of Twente, Enschede.

McKenney, S., Nieveen, N., & van den Akker, J. (2006). Design research from a curriculum perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 110-143). Francis & Taylor.

Mishra, P., Koehler, M., & Henriksen, D. (2011). The seven trans-disciplinary habits of mind: Extending the TPACK framework towards 21st century learning. *Educational Technology, 51*(2), 22-28.

Mislevy, R. J. (2011). *Evidence-centered design for simulation-based assessment.* Los Ángeles: The National Center for Research on Evaluation, Standards, and Student Testing (CRESST). University of California.

Morphew, V. N. (2012). *A constructivist approach to the national educational technology standards for teachers.* United States: International Society for Technology in Education (ISTE).

Nelson, B., Ketelhut, D. J., Clarke, J., Bowman, C., & Dede, C. (2005). Design-based research strategies for developing a scientific inquiry curriculum in a multi-user virtual environment. *Educational Technology, 45*(1), 21-27.

Nelson, B. C., Erlandson, B., & Denham, A. (2011). Global channels of evidence for learning and assessment in complex game environments. *British Journal of*

Educational Technology, 42(1), 88-100. doi:10.1111/j.1467-8535.2009.01016.x

Nokelainen, P. (2006). An empirical assessment of pedagogical usability criteria for digital learning material with elementary school students. *Educational Technology & Society, 9*(2), 178-197.

Oh, E., & French, R. (2004). Pre-service teachers perceptions of an introductory instructional technology course. *Electronic Journal for the Integration of Technology in Education*, *3*(1), 37-48.

Olasoji, R., & Henderson-Begg, S. (2010). Summative assessment in second life: A case study. *Journal of Virtual Worlds Research, 3*(3).

Plomp, T., & Nieveen, N. (2009). *An introduction to educational design research.* Enschede, the Netherlands: Netherlands Institute for curriculum development (SLO).

Quellmalz, E. S., Silberglitt, M. D., & Timms, M. J. (2011). *How can simulations be components of balanced state science assessment systems.* Policy Brief. San Francisco, CA: WestEd.

Quellmalz, E. S., Timms, M. J., Silberglitt, M. D., & Buckley, B. C. (2012). Science assessments for all: Integrating science simulations into balanced state science assessment systems. *Journal of Research in Science Teaching*, *49*(3), 363-393. doi:10.1002/tea.21005

Redecker, C. (2013). *The use of ICT for the assessment of key competences.* Sevilla: Joint Research Centre, Institute for Prospective Technological Studies. European Commission.

Redecker, C., Ala-Mutka, K., Bacigalupo, M., Ferrari, A., & Punie, Y. (2009). *Learning 2.0: The impact of web 2.0 innovations on education and training in Europe*. JCR Scientific and Technical Report.

Richey, R. C., & Klein, J. D. (2005). Developmental research methods: Creating knowledge from instructional design and development practice. *Journal of Computing in Higher Education, 16*(2), 23-38.

Rodríguez Espinar, S., & Prades, A. (2009). *Guía para la evaluación de competencias en el área de ciencias sociales*. Barcelona: AQU Catalunya.

Roelofs, E., & Sanders, P. (2007). Towards a framework for assessing teacher competence. *European Journal of Vocational Training*, *40*, 123-139.

Shute, V. J., Masduki, I., & Donmez, O. (2010). Conceptual framework for modeling, assessing, and supporting competencies within game environments. *Technology, Instruction, Cognition, and Learning, 8*(2), 137-161.

Silva, J. (2012). Estándares tic para la formación inicial docente: Una política pública en el contexto chileno. *Archivos Analíticos de Políticas Educativas, 20*(7), 1-36.

Sparrow, J. L., Blevins, S. J., & Brenner, A. M. (2011). Faculty development for and in virtual worlds. In R. Hinrichs & C. Wankel (Eds.), *Transforming virtual world learning* (pp. 47-65). United Kingdom: Emerald. doi:10.1108/S2044-9968(2011)0000004007

Stufflebeam, D. (2002). Foundational Model for 21st Century Program Evaluation. In D. Stufflebeam, G. Madaus, & T. Kellaghan (Eds.), *Evaluation models viewpoints on educational and human services evaluation* (pp. 33-84). New York: Kluwer Academic Publishers

Tessmer, M. (1993). *Planning and conducting formative evaluations: Improving the quality of education and training.* London: Kogan Page.

UNESCO. (2008). *Estándares de competencia en TIC para docentes.* Londres: Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura.

van den Akker, J. (2002). The potential of development research for improving the relation between curriculum research and curriculum development. In M. Rosenmund, A.V. Fries, & W. Heller (Eds.), *Comparing curriculum-making process* (pp. 1-12). Berlin: Peter Lang.

van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (2006). *Educational design research*. Francis & Taylor.

Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C.S., Estock, J. L., Orvis, K. L., & Conkey, C. (2009). Relationships between game attributes and learning outcomes: Review and research proposals. *Simulation Gaming*, 40(2), 217-266.

Woollard, J., & Wankel, C. (2011). Initial teacher training in a virtual world. In R.

Hinrichs (Ed.), *Transforming virtual world learning* (pp. 29-46). United Kingdom: Emerald. doi:10.1108/S2044- 9968(2011)0000004006

Zulkardi. (2002). *Developing a learning environment on realistic mathematics education for indonesian student teachers* (Doctoral dissertation). University of Twente, Enschede.

Figures

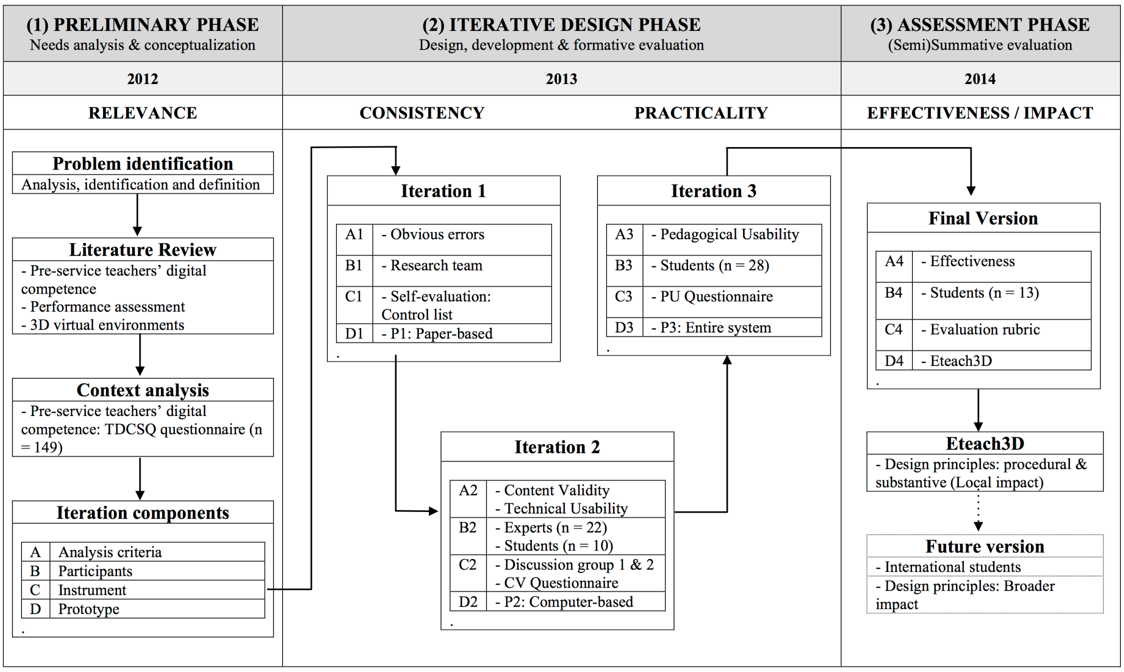

Figure 1. Process and structure of the study

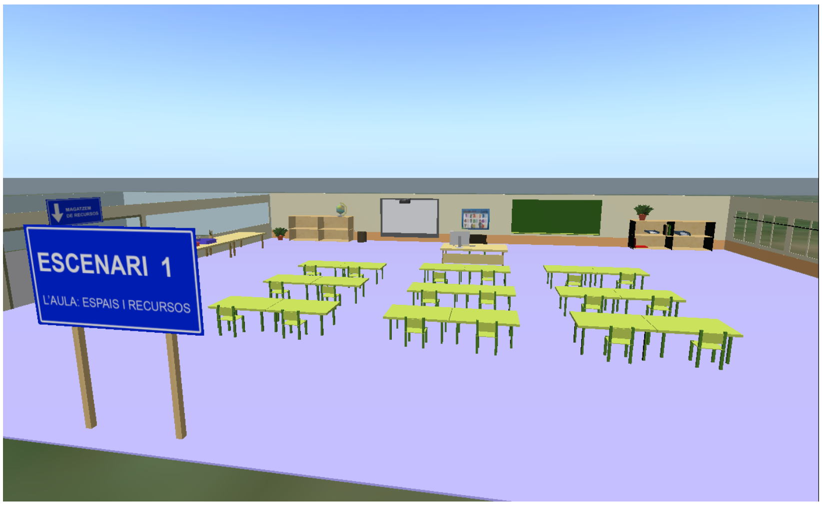

Figure 2. Primary school classroom of the second ETeach3D prototype

Figure 3. Primary school classroom of the third ETeach3D prototype

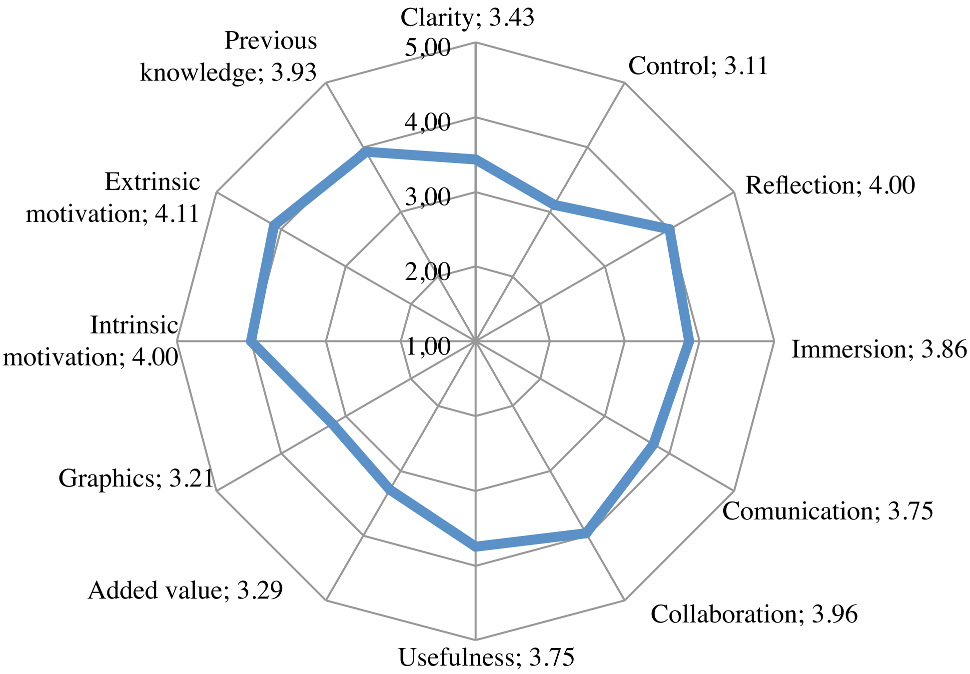

Figure 4. Average scores for the pedagogical usefulness of the environment (Scale: 1-5)

Figure 5. Images of students completing the evaluation activities for the final version of ETeach3D

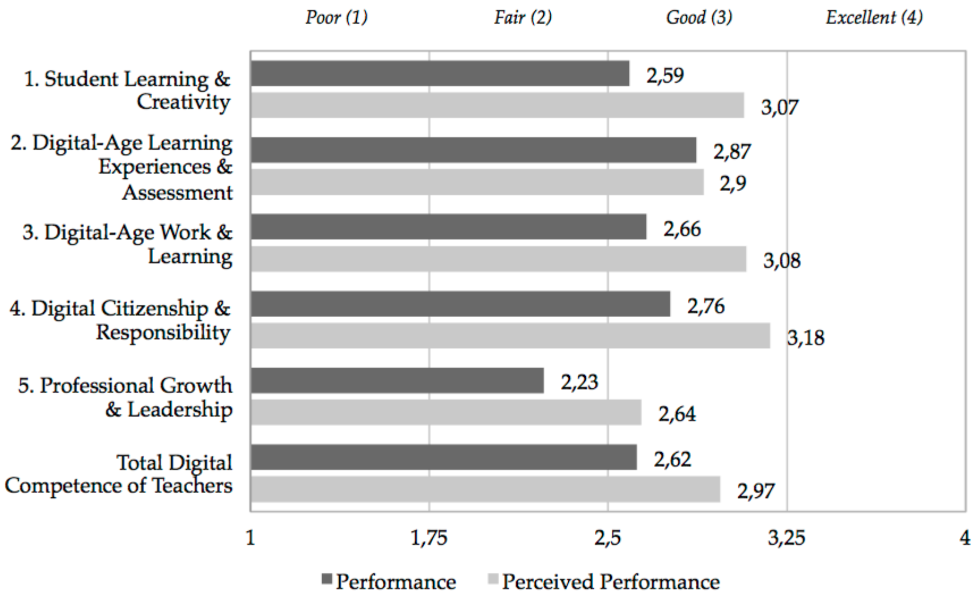

Figure 6. Comparison of the average scores for student performance and student self-perception (Scale: 1-4)

Tables

Table 1. Logic model for the ETeach3D

Inputs	**Processes**	**Outputs**
- Pre-service teachers - Digital competence: NETS-T ISTE - Performance assessment: ECD Model - MUVEs: Opensim and Sloodle	- Create 3D virtual scenes: (SC1) The classroom, (SC2) The workshop, (SC3) The staff room - Develop assessment activities - Design scoring procedures	- ETeach3D validation: content validity, technical and pedagogical usability. - Formative evaluation to improve the digital competence of pre-service teacher

Phase	**Step**	**Criteria**	**Instrument**	**Participants**
Preliminary	Context analysis	Self-perception	TDCSQ Questionnaire	Students (n = 149)
Iterative design	Iteration 1	Obvious errors	Control list	Research team
	Iteration 2	Technical Usability	Discussion Group 1	Experts (n = 6)
			Discussion Group 2	Students (n = 10)
		Content Validity	CV Questionnaire	Experts (n = 16)
	Iteration 3	Pedagogical Usefulness	PU Questionnaire	Students (n = 28)
Assessment	Final version	Effectiveness	Evaluation rubric	Students (n = 13)

Table 2. General process for data collection

Dimensions	**Average (SD)**	**Frequency (%)**							
		1	2	3	4	5	6	7	8
1. Student learning and creativity	6.04 (0.88)	–	–	1	5	17	41	33	3
2. Digital-age learning and assessment	6.11 (0.97)	–	–	2	3	18	34	38	5
3. Digital-age work and learning	6.17 (0.98)	–	–	1	3	18	37	33	8
4. Digital citizenship and responsibility	6.49 (0.90)	–	–	1	3	10	27	48	11
5. Professional growth and leadership	5.74 (1.22)	–	1	5	6	21.5	38	21.5	7
Total digital competence of teachers	6.11 (0.83)	–	–	1	4	14	44	36	1

Table 3. Descriptive statistics and frequencies of self-perceived teacher digital competence (Scale: 1-8)