

## **Virtual Reality and Education: Evaluating the Learning Experience**

*Maria Roussou  
makebelieve design & consulting, Athens - GR  
University College London. London, UK*

---

### **Introduction**

The majority of Virtual Reality (VR) applications developed today consists of research products that are either industrial prototypes created within very specific contexts or are used for presentation purposes. Despite the promise and the development activity of over two decades, there has been a considerable lack of real-world applications. The issues regarding the deployment of immersive VR in everyday work and education contexts have been discussed many times and continue to revolve around the familiar practical difficulties: setting up special and costly hardware within facilities that are not easily transportable, requiring special teams of developers and maintenance staff, but also providing the high-level tools that will support users in their complex tasks (Neale et al. 2002) and which can succeed in establishing a collaborative VR

work environment amongst individuals of different disciplines (Mackay & Fayard 1997).

One of the most important issues preventing the widespread use of immersive virtual reality has been the lack of evaluation efforts. As virtual environments (VEs) become more commonplace in practical situations, training, and education, there is growing concern about judging their outcomes.

As (Dede et al. 1996) stated in the mid 90's, "one of the biggest stumbling blocks in VR research right now is the lack of concrete data on the usefulness of VR". This is especially true with regards to virtual learning environments (VLEs), where relatively little principled empirical work has been carried out. In order to do this, effective evaluation methods need to be established to discover if conceptual learning takes place in VR (Whitelock et al. 1996).

Such evaluation methods were developed for the evaluation of the two different case studies involving educational virtual reality experiences for adults and children that are presented in this paper. In the first case of learning about archaeology and archaeological reconstruction, an evaluation study was performed in situ with adults and children who used the virtual environment during their museum visit. In the second case of understanding how to solve mathematical fraction problems, the evaluation study was held in a controlled laboratory setting.

### **Related Work on the Evaluation of Virtual Learning Environments**

A number of researchers have developed frameworks for structuring the evaluation of VEs (Gabbard et al. 1999; Hix et al. 1999; Bowman et al. 2002) However, most of these have been focused primarily on usability issues and usefulness for training and less on the efficacy of VEs for supporting learning in domains with high conceptual and social content. Gabbard et al. (1999), for example, propose a methodology for evaluating VE usability engineering, which involves sequentially performing user task analysis, expert guidelines-

based evaluation, formative user-centered evaluation, and summative comparative evaluations. In this process both quantitative and qualitative data are acquired, where qualitative data are typically in the form of critical incidents that occur while a learner performs task scenarios. A critical incident is usually a problem encountered by a user (such as an error, being unable to complete a task scenario, or user confusion) that noticeably affects task flow or task performance.

On the other hand, Marsh (Marsh et al. 2001) argue that standard human-computer usability evaluation methods, such as usability inspection, do not address the vicarious nature of activities performed within a 3D virtual environment through either a first person perspective, i.e. the point-of-view of the person immersed in the VE, or a third person perspective, i.e. a point-of-view from behind, over the shoulder or viewed from a fixed position, or that of an object or person representing the user.

Similarly, Cobb et al. (2002) note that the current immaturity of VEs in general and VLEs in particular has impacted on the types of evaluations that have been carried out. In their view,

## *Virtual Reality and Education: Evaluating the Learning Experience*

the most useful results found in virtual learning environment evaluations come from informal observed phenomena, not formal evaluations. They add, however, that it may be necessary for formal comparative studies to be carried out in order to show that the technology is educationally effective, an essential requirement if it is to be widely adopted in an educational setting.

Researchers concerned specifically with the evaluation of virtual learning environments have considered it important to investigate the educational efficacy of the medium in specific learning situations or broader learning domains, and develop new rubrics of educational efficacy that compare it to other approaches (Roussou et al. 1999). Dede believes that the efficacy of VR can be truly established only by rigorously comparing VR's benefits to traditional educational methods and only "through careful analysis that can accurately diagnose the weaknesses and limitations of the technology" (Dede et al. 1996). He and his colleagues organised the evaluation of their science learning environments around four basic aspects: usability, learnability, usability vs. learnability and educational utility. Neale et al. (1999) have

used Jonassen's constructivism principles (Jonassen 1988) as a framework for evaluating VLEs for students with disabilities and special educational needs. In their case, up to 8-minutes of verbal observation data per interaction were structured into a multiple activity chart in which student behaviour supporting each principle was coded when it occurred. Other attempts (Rose 1995; Salzman et al. 1999) at constructing methodologies and theoretical frameworks for evaluating learning activity within a VE have for the most part remained limited to particular applications and thus cannot be adopted to study specific aspects of the VR experience such as interactivity.

The question of whether VLEs require new and different evaluation methods beyond those in use by the HCI community or educational technology field, remains relatively unexplored. Many of the developed frameworks for evaluating learning follow a traditional approach, analogous to the standardized methods used to assess learning in formal educational contexts. This is not surprising as traditional educational assessment has proved to be remarkably resilient (Reeves & Okey 1996), despite growing criticism of its effectiveness in capturing what

really goes on in the learning process. However, the introduction of constructivist and situative perspectives of learning in educational practice has intensified the need to develop "authentic" evaluation techniques that are directly related to the learning approaches themselves. Hence, the increasing interest in alternative forms of evaluation is reflected in the proliferation of terms, such as authentic assessment, performance assessment, and portfolio assessment (Reeves & Okey 1996), that focus primarily on the process rather than just the product of learning. These methods may, in many cases, involve learners in the evaluation of their own learning, emphasizing common themes, such as problem solving and complex learning, which entail a wide range of responses and challenging tasks with multiple steps, time, and effort. Alternative assessment also affords the ability to include motivation as an important factor in the evaluation process. This is especially relevant to virtual learning environments which rely heavily on their motivational impact. The critics of alternative assessment, on the other hand, complain that it is time and labour intensive, and heterogeneous, as the outcome of the evaluation can vary wide-

ly in the specific knowledge domain being judged. It can also vary widely because individual students' performance varies. It relies on students' verbal and communication abilities and there is no easy comparison among students. Perhaps the most common critique states that alternative forms of evaluation can not be generalised to other contexts. However, cognitive psychologists as well as the education field, do not consider this a disadvantage, as they believe that the nature of knowledge itself is highly contextualised with limited generalisability (Brown et al. 1989).

Virtual learning environments are dynamic, contextually rich environments with a multitude of components that influence activity within them. The review of various methodologies and frameworks used to evaluate VEs has shown that none of the existing frameworks have been designed to capture the contextual, activity-based, and dynamically evolving nature of a virtual learning environment and of human behaviour with it or within it. Thus, a combination of evaluation methods and methodologies, presented below, has been considered as more pertinent in capturing the dynamics among all components.

## **Evaluation Methodology**

The evaluation methodology used for the case studies presented in this paper varies, depending on the purpose of the developed learning environment. In both cases, however, a choice was made to limit our testing to a small number of users and follow an in-depth qualitative approach, due in part to the nature of the projects but also to the fact that the participants in the evaluation sessions were children. Another reason for choosing a small number of users is the obvious practical difficulty of evaluations that are performed in situ - in our cases, getting museum visitors to agree in participating in experimentation which requires a significant investment in time and diversion from their planned visit schedule or getting parents to bring their children to a special VR laboratory. Finally, the highly experimental nature of some of the devices used, such as a robotic haptic interface, required significant effort to install and operate, which also hindered the evaluation process. For these reasons, case studies where small groups of users were studied in depth were considered more useful for gaining insights into the effectiveness and efficiency of the respective virtual learning environments.

For the first case study, our evaluation methodology draws from the structured framework proposed by Gabbard et al. (1999) and Schafer et al. (2002) for the design and evaluation of user activity in VEs. This includes the combination of user needs analysis, user task scenarios, usability evaluation and formative evaluation, and preliminary summative evaluation. The user needs analysis was carried out at the very beginning of the project and led to the definition of the user task scenarios that were used in the evaluation sessions. Even though the primary goal has been to evaluate learning effectiveness, usability evaluation formed a central tenet of the evaluation methodology, since all sessions involved observing the users of the VE in order to determine if the VE aided or hindered them in reaching their intended goals.

The detailed user requirements analysis with the different end-users, that is archaeologists, educators, and children (Roussou et al. 2004), confirmed the suitability of our choices and led to a detailed study of the existing workflow in these domains. Following the initial user needs analysis, we proceeded with the development of a complete VE for each case, which was continuously

informed by the participation of the end-users. Additionally, development and evaluation advanced together in order to determine the elements required to make the VEs useful in each learning context. The resulting VEs provide, in the one case a multi-modal learning environment for archaeologists, educators and students and, in the other case, an engaging play environment for children.

### **Methods**

The methods used in the evaluations included direct observation, questionnaires, and post-experiment interviews.

**Direct observation.** Participants performed the various tasks whilst being observed by a facilitator. Participants were encouraged to use a think-aloud protocol (Ericsson & Simon 1985) to explain what they are doing, to ask questions and to give information. Each participant was asked to concurrently verbalise her actions and thoughts whilst interacting in the virtual environment, while the facilitator used an interactive style, asking users to expand upon comments and activities. Sessions were also videotaped for further analysis. Direct, in situ, observation has been the primary method of

data collection for both studies presented in this paper.

**Questionnaires.** A usability questionnaire was developed to identify the user's perception of the effectiveness and efficiency of each VE and their level of satisfaction with the interaction. The usability questionnaire was constructed by merging a number of standard user satisfaction questionnaires, such as the approaches provided by Perlman (2004) and others (Davis 1989). The questionnaires include questions that require answers on a 1-7 Likert scale (Likert 1967). Additionally, pre-tests and post-tests were designed to identify participants' existing knowledge of the topic studied prior to entering the VE and the possible effect on participants after the experience.

**Interviews.** An interview following the experience was used to help identify the various issues that occurred during the experience and that could not be captured by the questionnaire. The interview was particularly important for understanding the issues involved in the in situ usage of the system, where the use of a questionnaire does not make sense. Given that the participants in these studies were young children or casual museum visitors, a combination of an

informal conversational and a semi-structured interview was chosen. Informal conversational, or unstructured, interviews are typically conducted in qualitative studies and allow the interviewer to ask the participant questions that emerge from the course of the discussion (Diamond 1999). The advantage of this informal kind of interviewing is that it increases the salience and relevance of questions, which can consequently be matched to individuals and circumstances. Another advantage of informal conversational interviews is their less threatening nature in comparison to more formal interviews (Diamond 1999), an important advantage when working with children.

The weakness is that different information is collected from different people with different questions, which can make data organisation and analysis difficult.

### **Apparatus**

The VR system used for testing was a CAVE-like display. The CAVE is a room-sized virtual reality system constructed of three translucent walls and a floor, onto which high-resolution computer-generated stereoscopic images are projected

(Cruz-Neira et al. 1993). The participant experiences the virtual world stereoscopically through a pair of active stereo glasses. In the first case study, a haptic interface was used to manipulate the virtual objects, i.e. the architectural elements of an ancient temple. In the second case study, a tracked position and orientation interaction device with a joystick and buttons was used to complete the virtual tasks assigned. The participant's head position and orientation in the CAVE is also tracked by a sensor, which is typically placed on the top edge of the stereo glasses. Due to the size of the glasses (which, unfortunately, are not designed for heads of different sizes, let alone children's heads), a more comfortable solution had to be devised (Fig. 36).

All sessions, in all studies, were videotaped. The camera was pointed toward the front wall of the CAVE, capturing each participant's back, the front screen, the floor and part of the side walls (Fig. 36). An external microphone connected to the video camera by a long cable was used to increase audio quality.

### **Case Study #1: VR in support of archaeological research and education**

The first case study concerns

the archaeological site of ancient Messene in Greece, specifically the study of the excavated Doric temple of the site. The temple is preserved in poor state, but there are a considerable number of architectural members found in the adjacent area, all of them well documented and interpreted by the Society for Messenian Studies, the archaeologists responsible for the excavation of the site.

We worked with the archaeologists of the Society in order to identify their needs and also to identify the users who would benefit from a VE that will visualize the process of an archaeological reconstruction (Fig. 37). As a result, restoration architects and archaeologists (especially archaeology students) were identified as the domain experts, who would use the VE as a tool for the exploration and validation of varied reconstruction hypotheses. To date, the tools available for this purpose are usually low-tech, low-accuracy models that cannot give the correct impression of scale and context. We also worked with the museum educators of the Foundation of the Hellenic World, a cultural heritage center in Athens with a CAVE®-like display open to the public, who expressed the need for a similar VE that could be used in the

context of a museum learning activity. In this case, adult and younger museum visitors, attracted by the interactive learning aspects of VR, would use a highly realistic interactive environment to learn more about history and archaeology.

The VE that was developed accurately represents a typical section of the temple, which the user must reconstruct. The user's activity in this VE resembles creative child's play with a construction kit: users of the environment must select the correct architectural members and position them appropriately. The process of virtually "building" parts of the temple provides the opportunity to actively experiment with different possibilities and solutions during the virtual reconstruction and to explore alternative scenarios. The environment also includes an instructional component, which provides novice users with information and the terminology used for each part of the reconstruction. Due to the tactile nature of the task, we decided to use a Haptic Interface, designed by PERCRO, as the main interface of interaction between the user and the VE.

#### Situated Use Sessions with Content Experts and Museum Visitors

The archaeological activity envi-



ronment was evaluated with expert and novice users in the context of a museum, with three different categories of users: adult novice users (museum visitors), young novice users (museum visitors between 9 and 14 years old), and adult domain experts (archaeologists and educators). All studies took place in the Foundation of the Hellenic World's cubic immersive display during or after normal museum hours. All novice users (adults and children) were family visitors who spent their day at the museum.

Overall, we ran complete sessions with a total of 14 adults, 7 of which were novice users and 7 content domain experts (Fig. 38) and with 7 children between 9 and 14 years of age (Fig. 39). In addition, we collected opinion questionnaires concerning the haptic interface from 25 more museum visitors after their experience with the environment, particularly the use of the haptic interface (Christou et al. 2006).

The instruments used were questionnaires and informal interviews. A usability and presence questionnaire was used after the experience for all users. Additionally, for the non-expert users, a pre-test questionnaire was used to test prior

knowledge and then a similar post-test questionnaire to see if there was a change in their knowledge, as a result of the virtual experience. We also collected general visitor opinions about the haptic interface with visitors of the museum who used it during normal museum hours.

### **Observations**

The evaluation of the virtual reconstruction case study with the content experts (archaeologists and educators) involved primarily the usability of the system and its potential as an educational work tool. The archaeologists we worked with and the majority of the archaeologists we evaluated the VE with, were very positive about the environment and its potential in educating restoration trainees, mostly because of its ability to present the content in a photorealistic and accurate manner and, most importantly, in the correct natural dimensions. However, most users pointed out that in order for the environment to be used in a real-world workspace (provided that all other practical issues were resolved), the representation of much more detail would be required, as well as the ability to simulate specific restoration techniques, such as filling in missing parts with plas-

ter or treating aging. Many comments concerning the potential of the haptic interface and suggestions for improvement were also collected.

The evaluation of the case study with novice users aimed at determining whether interaction within the VE helps the user to gain a better sense of the process of archaeological research and learn about the positions, dimensions, and inter-relationships between architectural members. The focus of the investigation has been on the potential for cognitive change, involving the measurement of the interaction effects on the user's understanding of the somewhat abstract concepts eluded by the task. However, an important aspect which cannot be separated from the evaluation of learning, especially when working with children, is the measure of affect (fun, engagement), as well as the potential pedagogical value of the system. Within this evaluation framework (cognitive, affective, pedagogical), we also looked at usability issues, involving mostly the learnability and ease of use of the system.

### **Case Study #2: VR in support of abstract learning for children**

The second case study concerns

a different learning domain, which was chosen to exploit the capabilities of the VR medium in visualizing abstract and difficult conceptual learning problems and providing feedback. This content domain could be no other than the field of mathematics education, implemented through an experimental VE in which children were asked to complete constructivist tasks that were designed as arithmetical fraction problems. The tasks designed for the virtual mathematics environment involved redesigning a virtual playground; specifically, modifying the areas that six of the main elements of the playground (swings, monkey bars, a slide, a roundabout, a crawl tunnel, and a sandpit) cover. Each virtual element covered an area which was colour-coded and represented by virtual blocks. This area representing each playground element was initially incorrect (either too big or too small) and had to be redesigned by the user, according to rules that require fractions calculations. Therefore, if the swings, for example, initially covered a  $3 \times 4 = 12$  block area in the virtual playground, the participant would be asked to find how many blocks to add or remove from the area in order to change its size. For the swings, the scenario required that the area be

increased by comparing two fractions (the fractions  $\frac{1}{3}$  and  $\frac{1}{4}$ ) and choosing the number that represents the larger amount. In this case, the fraction  $\frac{1}{3}$  which results in 4 blocks must be chosen and the 4 blocks must be added to the swings area, by picking blocks from the central pool and placing them on the 4 tiles of the virtual playground that need to be covered. Each study was conducted with one participant at a time lasting, on average, 90 minutes for each. The experimental methods included direct observation, interviews and pre- and post-test questionnaires, designed in collaboration with math teachers. Prior to the main activity, the participant was asked to fill out a questionnaire with math questions that are based on the fractions questions found in standardized tests (such as the Key Stage 2 SAT math test). A user profiling questionnaire was also given at this time. This included questions that attempted to draw a picture of the child's familiarity with computers, frequency of computer game play, and understanding of or prior experience with virtual reality.

### **Controlled evaluation with children**

Empirical work was carried out

with a total of 57 primary school students between the ages of 8 and 12, in different between-group experiments: an exploratory study, a pilot study, and a large-scale experiment. The exploratory study aimed at defining the evaluation methodology and framework for analysis. The pilot study, which was carried out a few months prior to the main experiment, aimed at improving the usability of the VE and helped in organising the overall process of the evaluation. The large-scale experiment, which took place in a controlled laboratory setting, involved a total of fifty ( $N=50$ ) children, 25 girls and 25 boys from different schools and socioeconomic backgrounds, who participated in one of three different conditions, two experimental VR conditions and a non-VR group. The instruments that were used to evaluate children's activity included direct observation, the conversational semi-structured interview, and written assessments of the topic, i.e. written questionnaires prior to and after the experimental tasks (Fig. 40).

The studies resulted in an enormous pool of data of multiple types, analysed both quantitatively and qualitatively. The quantitative analysis showed no meaningful association between

the different variables, such as gender, age, and condition, on student performance (measured through the pre- and post-tests). Therefore, we extracted specific examples, from the qualitative analysis, that provided us with interesting observations of student activity; instances of internal contradictions such as the ones that occurred during the analysis of the exploratory study. The pool of data was reduced -selected and condensed into a manageable form- by means of an inductive analysis, which produced central themes and patterns that emerged during this analysis.

### **Observations**

The analysis began by examining affective issues, i.e. participants' motivation and enjoyment, as potentially intertwined dimensions of activity in the VE that could relate to learning. Participant responses to the interviews and their observed behaviour indicated that the experience was highly enjoyed and intrinsically motivating. However, the methods to measure motivation and engagement and the constraints that they may carry with them were not adequate, thus concrete conclusions about how these elements relate to the unique properties

of VR and influence to learning in the VE could not be made.

Upon examination of the pre- and post-tests and the observation transcripts, some emerging themes were identified in terms of the learning content related to the tasks that proved to be difficult for many of the participants. The problems that more children seemed to have difficulty with were concentrated on issues such as fractions comparison or the confusion between the numerator and the denominator when asked to perform certain exercises. These problems were identified on the basis of the frequency with which they occurred in all three of the conditions of the study. What became evident during the analysis was that the visual and aural representational cues of the virtual environment, when coupled with the interactive VR system's feedback mechanisms, supported a certain type of activity and response on behalf of the participant which aided in problem-solving. The representational cues acted as visual forms of feedback for the participant, for example, judging if the area is a proper shape or guessing the number of blocks based on the available tiles and the surrounding space. Both cues and feedback created contradictions and then opportuni-

ties to predict contradictions. In this sense, the VR environment was very successful in supporting problem solving through a trial-and-error evolution strategy; consequently all participants in the study were able to complete all the tasks.

The role of interactivity in the VE proved to be central to this process of problem solving; interactivity was well suited in facilitating the operations level, i.e. aiding the participant in achieving the tasks by providing tools for successful planning and problem solving. The question posed by the evaluation study, however, was whether the interactive properties of a VE, e.g. system feedback, could enable the learner's transformation from conscious actions into operations, where planning and problem solving will have faded from the consciousness to give way to conceptual understanding. The analysis of the evaluation showed that learning environments should also involve guided interaction, permitting children to reflect on inconsistency and to change their conceptions (Roussou et al. 2006).

### **Conclusions**

The case studies presented above produced preliminary

observations which were derived from different types of experimental sessions and focus groups, with unavoidably small user sets. Although preliminary, with thorough analysis not yet completed, the results were rich because they provided insights and involved an in-depth observation of how actual non-IT expert users and children may be able to use Virtual Reality as a central tool in their learning. Conclusions at this point can only be general, the main one being that by using a learner-centered approach and a focused evaluation process, the development of VLEs can be tailored to the real needs of the end-users while the validity of the environments can be increased.

Many problems remain, of course. Firstly, these case studies are still far from proving that VR can be used in a real-world educational context with non-expert learners on a long-term basis. To date, there are examples of VR practice in industries, such as the automotive or oil and gas industry, where immersive systems are used in the workplace. However, even these workplaces still need to employ special laboratories and scientists in order to support the use of VR. When we talk about VR systems used in an educational

setting, in a leisure-based context, or for any other kind of widespread public use, we envision use that resembles (in terms of its simplicity and straightforwardness) that of a home or office PC. For this to happen, the practical difficulties of VE development, the issues of cost, distribution, space, and maintenance still hold and must be resolved, while the issue of evaluating effectiveness must become more systematic. If these issues are resolved, then virtual learning environments can become flexible and meaningful tools that can be used within real world learning environments and be of value to learners.

### **Acknowledgments**

The work described in the first case study was performed as part of CREATE, a 3-year RTD project funded by the 5th Framework Information Society Technologies (IST) Programme of the European Union (IST-2 0 0 1 - 3 4 2 3 1 ) , <http://www.cs.ucl.ac.uk/create/>. The work described in the second case study was performed at the University College London as part of the author's doctoral dissertation in Computer Science. Special thanks are due to all the technical staff and collaborators who

were instrumental in developing the systems and helped in the evaluation procedures.

### **References**

- BOWMAN, D.A.; GABBARD, J.L.; HIX, D. (2002). A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods. *PRES - ENCE: Teleoperators and Virtual Environments* 11: 404-424.
- BROWN, J.S.; COLLINS, A.; DUGUID, P. (1989). Situated Cognition. In Lawler & Yazdani (eds.), *Artificial Intelligence and Education*. Norwood, NJ: Ablex Publishing.
- BOWMAN, D.A.; GABBARD, J.L.; HIX, D. (2002). A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods. *PRES - ENCE: Teleoperators and Virtual Environments* 11: 404-424.
- BROWN, J.S.; COLLINS, A.; DUGUID, P. (1989). Situated Cognition. In Lawler & Yazdani (eds.), *Artificial Intelligence and Education*. Norwood, NJ: Ablex Publishing.
- CHRISTOU, C.; ANGUS, C.; LOSCOS, C.; DETTORI, A.; ROUSSOU, M. (2006). A Versatile Large-Scale Multimodal VR System for Cultural Heritage

*Virtual Reality and Education: Evaluating the Learning Experience*

Visualization. ACM Virtual Reality Software and Technology (VRST '06), Special session on VR in *Cultural Heritage, Education and Entertainment*. Limassol, Cyprus, ACM.

COBB, S.V.G.; BEARDON, L.; EASTGATE, R.; GLOVER, T.; KERR, S.; NEALE, H.; PARSONS, S.; BENFORD, S.; HOPKINS, E.; MITCHELL, P.; REYNARD, G.; WILSON, J.R. (2002). Applied Virtual Environments to support learning of Social Interaction Skills in users with Asperger's Syndrome. *Digital Creativity* 13: 11-22.

CRUZ-NEIRA, C.; SANDIN, D.J.; DEFANTI, T.A. (1993). Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE. *ACM SIGGRAPH: International Conference on Computer Graphics and Interactive Techniques*. ACM Press.

DAVIS, F.D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly* 13: 318-340.

DEDE, C. J.; SALZMAN, M.C.; LOFTIN, B.R. (1996). MaxwellWorld: Learning Complex Scientific Concepts Via Immersion in Virtual Reality.

*Second International Conference of the Learning Sciences*.

DIAMOND, J. (1999). *Practical Evaluation Guide: Tools for Museums & Other Informal Educational Settings*. AltaMira Press.

ERICSSON, K.A.; SIMON, H.A. (1985). *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA.: MIT Press.

GABBARD, J.L.; HIX, D.; SWANII, J.E. (1999). User-Centered Design and Evaluation of Virtual Environments. *IEEE Computer Graphics and Applications*.

HIX, D., SWANII, J.E.; GABBARD, J.L.; MCGEE, M.; DURBIN, J.; KING, T. (1999). User-Centered Design and Evaluation of a Real-Time Battlefield Visualization Virtual Environment. *IEEE Virtual Reality*.

JONASSEN, D.H. (1988). *Instructional Designs for Microcomputer Courseware*. Hillsdale, NJ: Lawrence Erlbaum.

LIKERT, R. (1967). *The Human Organization: Its Management and Value*. New York, NY, McGraw-Hill.

MACKAY, W.E.; FAYARD, A. L.

- (1997). *HCI, Natural Science and Design: A Framework for Triangulation Across Disciplines. Designing Interactive Systems (DIS)*. Amsterdam: ACM Press.
- MARSH, T.; WRIGHT, P.; SMITH, S. (2001). Evaluation for the Design of Experience in Virtual Environments: Modeling Breakdown of Interaction and Illusion. *Journal of CyberPsychology and Behavior* 4: 225-238.
- NEALE, H., BROWN, D.J.; COBB, S.V.G.; WILSON, J.R. (1999). Structured Evaluation of Virtual Environments for Special-Needs Education. *PRESENCE: Teleoperators and Virtual Environments* 8: 264-282.
- NEALE, H.; COBB, S.; WILSON, J.R. (2002). A Front-Ended Approach to the User-Centred Design of VEs. *IEEE Virtual Reality 2002*. Orlando, Florida: IEEE Computer Society.
- PERLMAN, G. (2004). *Web-Based User Interface Evaluation with Questionnaires*.
- REEVES, T.C.; OKEY, J.R. (1996). Alternative Assessment for Constructivist Learning Environments. In B.G.Wilson (ed.), *Constructivist Learning Environments: Case Studies in Instructional Design*. Englewood Cliffs, NJ, Educational Technology Publications.
- ROSE, H. (1995). *Assessing Learning in VR: Towards Developing a Paradigm Virtual Reality Roving Vehicles (VRRV) Project*. Human Interface Technology Laboratory, University of Washington.
- ROUSSOU, M.; JOHNSON, A.E.; MOHER, T.G.; LEIGH, J.; VASILAKIS, C.; BARNES, C. (1999). Learning and Building Together in an Immersive Virtual World. *PRESENCE: Teleoperators and Virtual Environments* 8: 247-263.
- ROUSSOU, M.; OLIVER, M.; SLATER, M. (2006). The Virtual Playground: an Educational Virtual Reality Environment for Evaluating Interactivity and Conceptual Learning. *Journal of Virtual Reality*.
- ROUSSOU, M.; SIDERIS, A.; LOSCOS, C.; DETTORI, A.; DRETTAKIS, G.; LOMBARDO, J.C.; COUDRET, F.; BIANCHI, C.; TECCHIA, F. (2004). *Requirements Analysis on Cultural Heritage - Education and Urban - Architectural Planning and Design Case Studies*. London, UK, University College London.
- SALZMAN, M.C.; DEDE, C.J.;



*Virtual Reality and Education: Evaluating the Learning Experience*

LOFTIN, B.R.; CHEN, J.A. (1999). A Model for Understanding How Virtual Reality Aids Complex Conceptual Learning. *PRESENCE: Teleoperators and Virtual Environments* 8: 293-316.

SCHAFFER, W.A.; BOWMAN, D.A.; CARROLL, J.M. (2002). *Map-Based Navigation in a Graphical MOO*. ACM Crossroads.

WHITELOCK, D.; BRNA, P.; HOLLAND, S. (1996). What is the Value of Virtual Reality for Conceptual Learning? Towards a Theoretical Framework. In P. Brna; A. Paiva; J.A. Self (eds.), *European Conference on Artificial Intelligence in Education*. Lisbon, Portugal. Lisbon: Colibri Editions.