Cybertherapy
Internet and Virtual Reality as Assessment and Rehabilitation Tools for Clinical Psychology and Neuroscience
G. Riva, C. Botella, P. Légeron and G. Optale (Eds.)
Amsterdam, IOS Press,© 2004, 2005, 2006

12 An Integrated Approach to the Ergonomic Analysis of VR in Psychotherapy: Panic Disorders, Agoraphobia and Eating Disorders

Carlo GALIMBERTI, Ph.D [1]; Gloria BELLONI, M.A. [1];
Alberto CATTANEO, M.A. [1]; Maddalena GRASSI, M.S.[1];
Valentina MANIAS, M.S. [1]; Luca MENTI, M.S. [1]

[1]*Catholic University, Milan, Italy*

Abstract. To face the aspects connected with VR environments' usability for psychotherapeutic applications means to dare a double challenge from a methodological point of view: from one side, the need to adapt and to integrate on a heuristic basis classic usability evaluation methods to specific artefacts such as 3D Virtual Environments for clinical applications; from the other hand, the problems arisen by integration of expert evaluation of VR environments user-based tests carried out in real context of use. The theoretical background of our analytical stance is based upon an ethnometodological approach, a perspective that gives evidence of how people, in specific social situations, are able to solve complex tasks producing shared meanings and achieving their goals during interaction. According to this perspective, the methodological objective consisted also in the identification of the usability requirements of the specific *community of practice.* The virtual environments considered were two of the four VR modules in the framework of the VEPSY project: Panic Disorders – Agoraphobia and Eating Disorders.

1. Introduction

This contribution presents the work and the results of the ergonomic evaluation and the tuning process of the virtual environments (VEs from now on) used for the Eating Disorders Treatment.

The study reported has its first reference in the *"Functional description of the ergonomic characteristics of the modules"* in which have been summarised and analysed the results of the functional ergonomic evaluations of the VEs that were used in the small clinical trials carried out on the four demonstration modules of the VEPSY project. The VEs have been developed during the first 15 months of the project according to the needs of the clinical protocols proposed by the clinical group and the hardware infrastructure and the software chosen, after having discussed different solutions, by the technological partners.

On the basis of tasks assigned by VEPSY Project to our research unit, we didn't oriented usability analysis on telepresence evaluation. This task had been accomplished by clinical units, while we focused on functional features of VEs. As we'll show later, our main goal was to get usability evaluation of processes of use of VEs as performed by 'real'

users in 'real contexts of use'. This is the reason why after the *functional analysis* of VEs carried out, we operated what we could call a sort of '*fine tuning'* of the VR scenarios.

Since the VEPSY environments are designed for clinical use, to fulfil our goals, further steps were necessary after the basic functional evaluation:

- To establish a minimum threshold of ergonomic acceptability to be used for every VR VEPSY module, on the basis of specific indicators identified and reported in the Guidelines prepared at the beginning of the project

- To develop a new method of ergonomic and usability evaluation taking into account the requirements needed by the specific typologies of end users:
  - Psychotherapists
  - Patients affected by specific psychopathologies

- To integrate the results on the basis of the observations emerged after *Large Clinical Trials*: this implied a direct confrontation and a synergic collaboration with the clinical group

The basic assumption is that VEs should enhance the physical, cognitive and perceptual capabilities, allowing users to do things that are not possible – or 'not always possible', as for specific pathologies – in the real world. As stated by Mantovani [1], difference is not made by the realism of the experience but by the therapists' and patients' shared experiences as a whole, defining a clinical protocol to make use of the interaction and of the tools. As stated by Sutcliff, Gault and Shin [2], quality assessment of VEs has tended to focus on assessment of presence, trying to evaluate "how real or natural the user's experience was when immersed in the environment. Presence has been evaluated by questionnaires which ask users to rate various qualities of the VR environments ranging from perception of «being there» (Slater et al.) [3], to more detailed inventories ranking controls, feedback, perception of realism and user engagement". As we have already stated, telepresence was not a variable we had to evaluate. In addition to this, we didn't paid much attention to it because we share Sutcliff position on presence's measures when he says that they "can benchmark VE designs in terms of their realism and overall user experience" but "they do not help to diagnose design flaws for formative evaluation" [2]. For these reasons, in this paper we propose a set of heuristics and an expert evaluation method that, stemming from the evaluation of user interface, has been extended to VEs, with the goal to evaluate them in their real context of use. This assumption made the attention to be focused on two specific aspects:

- The consideration of areas of VR not yet completely explored from the usability point of view. For instance, the concepts of spatiality and representation (construction of meaning)

- The transfer of the results on the basis of a user-centred approach strongly applied to the real context

1.1 Scheduling of the ergonomic evaluation process

Our research unit conducted the ergonomic evaluation of two of the four VEPSY modules in 3 Phases:

1 **Panic Disorders and Agoraphobia modules**: in ***Phase 1,*** *guidelines on heuristic basis* were prepared in order to have an effective evaluation tool. Afterwards, usability test (observations) on generic users were carried out
2 **Eating Disorders modules**: in ***Phase 2,*** *basic functional requirements were verified*, referring to the results obtained in *Phase 1*. Then, usability tests (observations) were carried out on a different sample in comparison with the *Phase 1*, considering psychologists and non psychologists.
3 **Eating Disorders modules**: in ***Phase 3***, *semi-structured interviews* were carried out on psychotherapists involved in the clinical trials of the modules considered.

1.2 A psychosocial integrated approach

The possibility to settle a research implant taking into account the cultural using context, the bargaining character of interaction and of its intrinsic 'opacity', represents the main methodological objective of this study. In particular, an ethnomethodological perspective is adopted.

As stated by Zucchermaglio [4], "Ethnography is one of the most adequate methods to enter communities by interpreting those meanings which are relevant for members in building up and interpreting the social world, looking for them in the discoursive interactions and in the public intersubjectively accessible behaviours. The validity of the ethnographic research is not to be found in the objectivity of the description, but rather in the level of authenticity, plausibility and reliability provided by the descriptions also to the observed subjects (...). We underline, for the comprehension of social situations, the importance of the *categories of meaning* performatively used by people involved in those specific situations".

1.3 The classic usability perspective

As emphasised by Cantamesse and Menti [5], the nowadays-considered usability is strictly connected with the evolution of ergonomics, term defining "a new method to study and to solve the relationship between the man and the working environment". The application field of ergonomics is rather wide: from the design of common use items to the user-computer and user-computer-user interaction with reference to the technologies. The viewpoint is a multidisciplinary one. Psychology and especially Social Psychology played a significant role in the evolution of the ergonomics towards a preventive rather than a corrective function. With the introduction of the prevention ergonomics the division between human error and machine as the relational concept has been adopted, centred on the relationship among persons, environment and tools used. From this moment on the artefact is considered as "an experience transformer": what was once defined as a task has become just a part of a wider scenario. Informatic artefacts became first of all "communication tools" and then "integrating part of the communication". In this development process, the concept of *usability* replaces the one of *reliability*. Since 1998 the definitions of usability suggested by the International Organisation for Standardisation (ISO 9242) evolved with stronger attention to the importance of the user even considering different experimental setting for the different subsystems: effectiveness, efficiency, satisfaction, understandability, learnability, operability [6]. Jordan [7] suggests a usability model including intuitiveness, learnability, expert user performance, system's potential, and re-usability.

As to the usability research methods now in use, Pedon makes a distinction between experimental and non experimental methods: the application choice is made upon higher or lower need to control dependent, independent variables and sample selection [8].

Techniques can be quantitative or qualitative. The quantitative ones give important information on the artefact failures or on the users' behaviour, but are not explanatory of the causes underlying those specific behaviours or of the services required. Qualitative techniques allow to analyse the whole interaction context including needs, aims and mental models. The *expert assessment*, aided by guidelines and checklists according to the heuristic evaluation, and the *user-based testing* [20], implemented with other techniques and tools such as interviews and focus groups, play the most relevant role.

Till now, despite intense and wide-spread research in both usability and VEs, there is no evidence that improvements in the first field could be applied to VEs' evaluation. As far as we know, VEs' new technology has not yet been closely and sufficiently coupled with the important characteristic of usability [9]. Too often we see usability evaluation methods of interactive computer applications, whose limitations are well-known, are adopted for evaluating VEs. And this is the reason why we think we need to develop usability evaluation methods and criteria *specifically* for VEs. As Sutcliffe has stated "few evaluation methods have been proposed for assessing the usability of VEs, although field studies of VR designers have demonstrated the need for HCI knowledge and methods" [2].Obviously, the point has been discussed by several papers: Gabbard and Hix, for instance, have tried to highlight usability problems associated with the use of VEs [10].

Bowman and Hodges, among others, have shown that the designers of VE systems cannot rely only on the methods developed for standard graphical user interfaces (GUIs) because of the fact that their interaction styles are totally different from standard user interfaces [11]; Johnson has tried to exploit design principles to generate heuristics to evaluate desktop VR applications [12] while Kalawsky adapted checklist evaluation methods, based on Nielsen's heuristics (1994), for VEs [13]. Generally speaking, most studies repertoried by Sutcliffe "have followed observation and expert interpretation of users' errors [14] or experimental studies reporting performance data and problems in a range of VE technology [15]". Nevertheless, we think that Gabbard statement [16] that "researchers interested in VE usability are left to performing ad-hoc assessment or in-house evaluations with little or no scientific basis for their approach" couldn't be kept today. In fact, we think that – thanks to recent developments in ergonomics studies – we now have almost all that is needed to develop a "a method to guide the usability evaluation of virtual environments".

In this paper we try to show how is possible to develop a psychosocial model of usability for VEs based on evidence of how people, in specific social situations, are able to solve complex tasks and produce shared meanings while completing these tasks in VEs in order to make their actions understandable and successful.

1.4 The analysis procedure: towards an 'ecological context of use'

To fulfil this goal we had to shift our attention from VEs in themselves to the relationships between users and VEs, focusing on how these relationships took shape in their real context of use. To approach the most ecological context of use, we turned to the LPP model (Legitimate Peripheral Participation): "Such a model considers the knowledge acquisition in progression terms – from the periphery to the centre - in the participation activities of the *communities of practice"* [17].

The study has thus been structured in three phases, each of them featured by:

- specific aims
- specific objects (i.e. two VEs typologies for different psychological disorders)
- samples reflecting aspecific, specific and professional ethnomethods
- generic, finalised and lived experience analysis contexts

Table 1. The research framework

Phase	Specific goals	Specific objects	Samples	Results	Analysis context
1	Functional characteristics	Panic Disorders Modules	**A** Generic users n=33	Aspecific Ethnomethods	Generic contextualization
2	Fine tuning	Eating Disorders Modules	**B** Psychologists /non psychologists n= 16	Specific Ethnomethods	Finalized contextualization
3	Integration	Panic Disorders Eating Disorders Modules	**C** Psychotherapists involved in clinical trials n=4	Professional Ethnomethods	Lived experience

2. Methods

The approach applied concentrates on identifying usability defects. Defects can be identified by expert assessment, aided by guidelines and checklists (heuristic evaluation) or by user-based testing (Nielsen, 1994) [18]. Moreover, classic usability methods have been integrated with specific ethnomethodological tools [19].

2.1 Main usability evaluation methods applied

Functional Analysis aided by expert heuristic evaluation**:** it is a type of analytical evaluation in which an expert in user interaction design assesses a particular user interface by determining which usability design guidelines it violates and supports. In this specific case this aspect was particularly challenging because, at present, there are very few guidelines specific to VR user interfaces.

To overcome this problem, the research unit prepared specific *Guidelines* to be applied to the expert evaluation and to the user-based tests. They were also designed with the aim to represent a flexible tool allowing further adaptation required by the different functional characteristics of the VR scenarios: for example, in Panic Disorders and Agoraphobia VEs, the evaluation of the commands toolbar was necessary, while in the Eating Disorders environments it was not present.

The heuristics used in this study are derived from Nielsen [20] and represent an attempt to consider areas of VR not yet completely explored from the usability point of view (i.e. spatiality: the use of spatial dimensions; virtuality: the sense of telepresence; representation: the construction of meaning) together with 4 indicators used to evaluate as many critical areas:

- Navigability
- Expected utility
- Communicative efficacy
- Graphic appeal

The *Guidelines* were intended to be:

- a reference tool for usability evaluation during the different phases of the modules implementation and tuning
- a supporting tool in identifying intervention priorities with particular reference to the interfaces
- a supporting tool to integrate modules in a coherent package
- a supporting tool for the definition of the characteristics of the final training activities
- a flexible tool for next evaluation activities concerning the final product

All along the study, we had to evaluate consistency and validity of the various heuristics we found out during the first analysis, as they resulted from the continuous regulation between users' performance's needs and minimum usability condition of each VEs. [21, 22].

Functional analysis performed in phase 2, for instance, is based on the reconsideration of the listed below critical points revealed by the consideration of VEs' 'navigability' and 'graphic appeal' in phase 1:

- movement control z axis
- collision control
- world's limits
- possibility of interaction
- anchorage of the objects

User-based test: it is a type of observational assessment with users that begins in the earliest phases of user interaction design and continues throughout the entire life cycle of the designing process. Qualitative data are usually in the form of critical incidents occurring while a user is performing task scenarios: they are events affecting user performance either in a positive or in a negative sense. Main focus of the research was to integrate this method with specific tool and by using different modalities in order to investigate the concept of VR usability in the real context of use. This goal was achieved gradually.

2.2 Specific tools

Classic usability evaluation methods have been supported by two specific ethnomethodological tools: *micro-narration* and *interviews*, used respectively in phase 2 and in phase 3. In the frame of the paradigmatic change under consideration, a relevant role is played by the narrative concept of knowledge and culture.

Narration can be considered both an adequate tool to recover shared practices, in particular, through recollection, and also a useful tool to create a group culture, that is to suggest a repertory of meanings determining what is important to observe in connection with consolidated habits. In the different phases of the analysis, micro-narrations and interviews were presented to subjects in order to recover information related to the co-interaction with the artefact and with the co-construction of meanings in a specific professional community.

Micro-narrations: users are supplied with specific information helping them to interact "*as if*" they were in the real context. For example: basic information about the specific VR protocol and about the therapeutic setting were given to the psychologists tested. They were informed that "purpose of the environments is not the creation of a perfect reproduction of the real world: patients and therapists involved are aware of the fact that

the effectiveness of the tool for the patient does not depend on the perfect accuracy of some specific elements but on the feeling of presence perceived that could be very different from the one of a person without pathologies".

Non psychologists users were asked not to consider VEs as videogames. They were explained the potential applications of the environments considered.

*Semi-structured interviews***:** in-depth interviews with the clinical group (psychotherapists involved in the clinical trials) were carried out in order to move towards an ecological context of use. In the specific case, the investigation was focused on 4 main areas with reference to ergonomic aspects:

1 In context use of the VEs for psychotherapeutic sessions
2 Expectations of the therapeutic protocol
3 Usability (local interaction with the artefact, interpretation of the situation and context definition)
4 Towards a culture of use (possible future application of the VEPSY modules; critical aspects for training activities etc.)

3. The research

System characteristics:
Pentium III 750 Mhz - Windows 2000 SP 1- 128 Mb RAM – Hardware 13 Gb – Flatscreen 14" – Graphic board: Matrox G450 Dualhead
Headset: Sony Glasstron 800X600 – 75 Hz / Tracker: Intersense Intertracks 30

Phase 1
Modules: Panic Disorders - Agoraphobia

The *usability tests* were conducted on 33 subjects (16 immersive mode, 17 non immersive mode) belonging to both sexes, aged between 20 and 26. When building the sample the participants' degree of skill in navigating in VEs was not considered as independent variable.
Observation method: Thinking aloud (16 subjects) and Aided Interaction (17 subjects).
Time per session: 45 minutes
A *focus group* (5 subjects) was carried out after the tests.

Phase 2
Modules: Eating Disorders

The *Expert Functional Desk Analysis* (verification of the functional requirements) was conducted both in immersive and in non immersive mode on the basis of the results

Table 2: Panic Disorders and Agoraphobia usability test – Sample A

	IMMERSIVE MODE		NON IMMERSIVE MODE	
	Thinking aloud	Guided interaction	Thinking aloud	Guided interaction
n	8	8	8	9

Total = **33**

Table 3. Eating Disorders usability test – Sample B

Immersive								Non Immersive							
Psychologists				Non Psychologists				Psychologists				Non Psychologists			
18-40		40 - *		18-40		40 - *		18-40		40 - *		18-40		40 - *	
M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Total = **16**

Table 4. Semi-structured interviews– Sample C

	Panic Disorders – Agoraphobia	Eating Disorders
Psychotherapists involved in clinical trials	1	3

emerged during *phase 1*. This model of analysis granted a considerable saving of time and of resources since it allowed an easier categorisation of the critical elements deserving special attention. Moreover, the utility increases, considering that the new module is made up by a number of VEs definitely higher than the one previously analysed (*Panic Disorders' - Agoraphobia modules*).

The *usability test* was conducted on 16 subjects (8 immersive mode and 8 non immersive mode) belonging to both sexes aged between 18 and 40 and over 40.
Time per session: 45-55 minutes

The users' sample consisted of psychologists and non psychologists. Specific ***micro-narrations*** for **psychologists and non psychologists** have been used in order to create a more ecological context of analysis.

Phase 3
Modules: Panic Disorders – Agoraphobia / Eating Disorders

Semi-structured interviews were conducted with the 4 psychotherapists involved in the clinical trial: 1 for Panic Disorders – Agoraphobia; 3 for Eating Disorders.

4. Results

4.1 Phase 1

Usability test - Main function of this analysis was the identification of variables useful for a rough elaboration of the usability components deserving major attention in the *phase 2* (Fine Tuning). The results of *phase 2* were then used to tune in the modules with the real context of use and with the specific requirements of the *communities of practice* to which they are destined to.

Specific objects in phase 1 were the 4 VR scenarios composing the Panic Disorders' and Agoraphobia modules (Lift, Underground, Plaza, Supermarket). The interaction in each of the 4 environments was analysed in detail according to 3 dimensions:

- Navigability
- Expected utility

- Contents completeness/graphic appeal

Fifteen days after the usability tests, a focus group was carried out in order to validate what was recorded through the observations. A synthesis of the emerged results follows, stressing out basic observations concerning each of the 4 considered scenarios. The critical surfaced topics have been used to focus the dimensions deserving special attention during *phase 2* and to carry out basic modifications in the course of the designing process, taking into account that the methodological objective consists also – as already stated – in the identification of the requirements of the specific *community of practice* the modules are destined to and not only the needs of a generic sample of users. This is the reason why, for example, some elements identified or perceived in this first phase as 'limits' to the interaction could even represent strong points when considering the specific ethnomethod.

In particular, it seems important to consider some factors influencing and orienting the experience and assessment that subjects gave of it as, for example, expectations about the experiment and about the navigation experience in a VE. In *phase 2* a specific pre-test and post- test interview about expectations was included to better investigate this specific aspect.

Here follows an example of the results and basic modifications suggested in order to improve the environments' usability.

Evaluation - rating: 1-5 scale

4.2 Phase 2

In the phase 2, on the basis of the outcomes of the previous phase, a preliminary functional analysis was carried out, considering each scenario composing the Eating Disorders' VR modules.

Table 5. Example of a usability test's outcome – phase 1

Lift

Indicators	**Test results**	**General evaluation ad comments**
Navigability	▪ The environment is limited but there are good possibilities of exploration ▪ Rotation effect on 'x' axis was ok ▪ Opening the lift door is not intuitive ▪ 'Reset world' and 'level the world' are often used	++
Expected utility	The possibility of being able to move with the lift was disregarded due to the absence of the push button panel inside the elevator	++
Graphic completeness/appeal	▪ Different levels of graphic realism inside and outside the elevator ▪ Necessity to add the 'push button panel' inside the elevator	++ add the 'push button panel' inside the elevator graphic definition to be improved

As underlined by Sutcliffe [23][1] "since VEs are constructed to represent the real world, user tasks should ideally mirror real world actions; however, in practice limitations of technology mean that some compromises have to be accepted. above all considering real requirements of specific users."
In particular, evaluators focused their attention on the presence or absence of features in the following categories.

Lack of haptic feedback. As known, virtual prototypes have no haptic feedback (sense of touch) so the user's presence can pass through representations of solid objects. Following Sutcliffe, we could say that "to mitigate the lack of haptic feedback many applications use visual feedback with collision detection algorithms to prompt users when objects are selectable or have been selected. Problems caused by absence of haptic feedback may be observed with complex manipulations and physical tasks. These problems can be avoided by designing augmented reality in which interactive surfaces are modelled as physical mock-ups, but in many VEs this is too expensive".

Realistic graphics. Sometimes VEs are not able to represent the prototype, and this because most applications are not rendered in photorealistic detail. As Gabbard says, there is some evidence suggesting that people can perform tasks naturally even in absence of detailed visual cues and representations [24]. Nevertheless, we must admit that graphical detail is important for information displays and for tasks when the system environment is visually complex. The same could be said for audio input.

Compatibility with the user's task and domain. The VEs and behaviour of objects should be as similar as possible to the user's expectation of real world objects, of their behaviour, and affordances for task action.

Faithful viewpoints. "The visual representation of the virtual world should map to the user's normal perception, and the viewpoint change by head movement should be rendered without delay".

Navigation and orientation support. "The users should always be able to find where they are in the VE and return to known, preset positions. Unnatural actions such as fly-through surfaces may help but these have to be judged in a trade-off with naturalness".

Clear entry and exit points. "How to enter and exit from a virtual world should be clearly communicated".
Number of tester: 2

General considerations:

Menu/commands:

- In full screen modality it is not possible to access the menu
- Changing environments: by pressing the key "Esc" and by dragging the icon into the browser's window
- The commands' menu is complete. More details are needed about the functions available
- Movements: on the keyboard only go-forward and go-backward movements are allowed

[1] Since we found our results very keen to what Sutcliffe said, we made a large use of his formulation of Nielsen heuristics

Scenarios:

- Movements along the 'z axis' are not continuous. In particular: lowering the height level in the 'go-backward' movement.
- Collision control: with peripheral walls only
- No interaction with objects with the exception of some specific situations. As for the doors, the lack of interaction (possibility to open the door) sometimes causes uselessly navigation
- The world's limits control has not always been implemented

Evaluation - rating: 1-5 scale

Table 6. Example of a functional desk analysis' outcome

Scenario	**Control on 'z' axis**	**Collision control**	**World's limits**	**Interaction**	**Details' quality**	**Anchorage**
Kitchen	+++	no	no	yes	++++	no
Comments: interaction with refrigerator's doors, cupboards, oven by double clicking the 'enter' key (on the keyboard); it is possible to virtually eat the food by double clicking the 'enter' key (on the keyboard): food appears on the table. Graphic details of the objects inside the refrigerator are not perfectly refined. Immersive mode: movements by 'fits and starts'						
Bedroom	++	yes	yes	no	+++	no
Comments: the predominance of the blue colour gives a sense of oppression – no risk of losing the sense of direction also because the room is small. Graphic improvements suggested: curtains, light sources						
Bathroom	++	no	no	no	+++	no
Comments: Texturing can be improved – Movements are rapid – The quality of the sauna's graphic details can be improved – Difficulties in navigation related to the architectural characteristics (staggered walls, partitions) – The scenario is realistic						
Shopping Arcade	++	no	no	no	++	no
Comments: There is no graphic continuity among the shop's walls – Graphics to be improved in the definition of the ceiling – Texturing quality level not homogeneous – Graphic definition of the objects is not homogeneous. Perspective defects with the walls' angles						
Psychologist's room	+++	no	yes	no	+++	no
Comments: movements by 'fits and starts' – texturing and graphic definition not homogeneous – No risk of losing the sense of direction also because the room is tight						
Weighing machine	+++ limited movements	no	yes	yes	+++	no
Comments: Full-screen modality: there are no instructions to enter weight data – The button 'passa' doesn't give feedback						

Sitting room	++	yes (table)	yes	no	++++	no
Comments: difficulties in 'go-backward' movement						
Safe place	+++	no	no	no	++	no
BIVRS	++	no	no	yes	+++	no
Comments: difficulties with the sense of direction due to poor graphic details in the entrance hall						
9 doors' room	+++	yes	no	yes	++	no
Supermarket	+++	yes (shelves)	yes	no	++++	no
Comments: Graphic defect: wrong specular rotation of some images on the shelf						
Gymnasium	+++	no	no	no	++	no
Comments: Graphic not well refined – the scenario is realistic						
Pub	++	no	no	yes	+++	no
Comments: Interaction with the menu on the table. It is possible to 'open' the menu by double clicking the enter key on the keyboard. Chosen food appear on the table – Graphic details can be improved (block out)						
Clothing shop	++	no	no	no	+++	No
Comments: Graphic definition of the dressing room can be improved – the scenario is realistic						
Restaurant	++	no	no	yes	+++	No
Comments: Interaction with the menu on the table. It is possible to 'open' the menu by double clicking the enter key on the keyboard. Chosen food appear on the table – Graphic details can be improved (block out)						
Beach and Pool	+++	no	no	no	++++	No
Comments: Movements are very slow – Graphic details can be improved (block out) – the scenario is very realistic						

Usability test – In the second test session, some indicators were added to the ones previously used in order to evaluate specific cognitive aspects.

Table 7. Example of a usability test's outcome – phase 2

	Psychologists	**Non Psychologists**
Usability	Navigability	Navigability
	Communicative efficacy	Communicative efficacy
	Content completeness Graphic appeal	Content completeness Graphic appeal
	Expectations	
Cognitive aspects	Attitudes	
	Representations	
	Product evaluation	

Kitchen

Navigability	**Communicative efficacy**	**Content completeness** **Graphic appeal**
▪ Indecision about how to start exploration. This seems to be due to the plenty of stimuli: "I can see the oven...I'm going to see if it works...no I want to see the view out there, before" ▪ Tendency to go backward to have a better perspective ▪ Frequent loss of spatial co-ordinates (above all, along the Y axis) +++	▪ Even with aided interaction, the subject has difficulties in interacting with the objects ▪ Difficulty in positioning correctly in front of the objects to try the interaction (food, cupboards) ▪ The possibility of interaction is not perceived: there are no attempts to interact +++	▪ The scenario is judged complete and detailed ▪ Some specific elements like food or pots draw the attention: "The apples seem to follow me, I can always see them" ▪ In general, some elements are better refined than others from a graphical point of view ▪ Some expert users notice the absence of light sources ++++
Immersive mode /non immersive mode : ▪ Immersive mode: difficulties in keeping the same height level in moving forward and backward ▪ Tendency to lower the head ▪ The restart command is often used		

Evaluation of cognitive aspects

Psychologists

Expectations

- Among expected advantages:
  - Rapidity of access to the patient's dysfunctional representation
  - Facilities for the therapy: i.e. the outdistancing from the dysfunctional representation
  - Possibility to modulate some elements in the scenario according to the level of the pathology and with reference to the established therapeutic path
  - Opportunity to reduce the patient's discomfort

- *Interaction modality*: expectations are oriented towards the possibility to find in the VR situations tasks to perform

Attitudes

- During the interaction most of the psychologists pay attention to the aspects connected with **the emotional assimilability to the painful situations experimented by patients rather than to graphic/technical details**
- The typology of the environments is judged effective. The same for the disturbing effect caused by people represented in the scenarios (social element)
- Some psychologists express perplexity about possible reactions of anxious patients to sudden black-out of the system, loss of directions or excessive slowness in some environments

Representations

- Many of the considerations are referred to the specific theoretical approach of the psychologists interviewed
- The specificity degree of the different scenarios is evaluated very effective with reference to the specific pathologies considered
- Generally speaking, users expect an higher level of interaction

Product evaluation

- Good. The basic concept of the environments is evaluated applicable
- Navigation is not considered intuitive at all but anyway easy upon correct instructions
- Users underline the possibility to improve data collection and therapeutic process monitoring offered by this kind of product
- Modulability and diagnostic applications are considered the strong points
- Users stress out the necessity to make this kind of technology easy accessible

4.3.1 Phase 3

Semi –structured interviews: data collected in phase 1 and 2 were integrated with data collected through semi-structured interviews carried out with the clinical group. Critical usability indicators were identified and a final evaluation was made with strong reference to applicability and use in *real context*.

General information about the use of the scenarios

Patients' selection criteria

Patients were selected according with their pathologies. Female patients aged between 18 and 50 were considered as they represent the most frequent population in the hospital where the clinical group is operative. They could not have psychiatric pathologies.

They had to possess a PC at home to participate in the following telemedicine phase by using the chat modality; for the control group the distance communication tool was the telephone.

Use of the scenarios

The standard protocol foreseeing a sequence of immersive experiences for each pathology was followed. Each researcher carried out the clinical trial with patients affected by all the pathologies considered.

Protocol application

The protocol foresees a first contact with the VEs through the 'safe place' (psychologist's room). This scenario is used during initial session as training module for patients (commands' use and navigation). In the following sessions, patients start with most restricted rooms/environments towards the widest ones.

Expectations of the therapeutic protocol

Expectations were satisfied by the clinical trials that demonstrated the validity of the tool in the framework of a standard therapeutic protocol. Doubts about the acceptation level by patients were disconfirmed: they showed a high degree of involvement, subsequently confirmed in the post-treatment phase.

Protocol main differences in the traditional therapy

VR offers the possibility to monitor the patient's reactions in real time and to measure variations of physiological and psychological reactions after a cognitive reconstruction phase. *"...the patient does not have to imagine the real world: the vividness is the same and stimuli can be controlled."*

Traditional therapy foresees a cognitive reconstruction of dysfunctional thinking processes by using imagination. With the VR therapy it is possible to speed up these processes, because the observation of the patient's experience gives the therapist the possibility to interact actively with the patient. He can describe the mechanism regulating the pathologic behaviour bringing it on a 'meta' level where the therapist can operate "...once I was in the supermarket, I remembered what I saw in my eyeglasses (i.e. headset) and everything my therapist told me"

Usability evaluation

Critical aspects identified in *phase 2* were rated on a 1-5 scale considering the interaction modalities in real context of use.

Critical aspects: Navigability: Axis control – Collision control – World's limits - Anchorage

During the VR therapy the patient is not free to move around: he is accompanied by the therapist that guides him on a specific pathway and operates in order to facilitate the interaction of the patient inside the VR situation. There is no possibility to lose the sense of direction or to try the 'space suspension' experience that could interrupt the continuity of the immersive experience.

Orientation – Adaptation to the VR experience: small rooms (i.e. psychologist's room) have the function of allowing patients **to adapt themselves to the VR experience**. They are also used because of their limits in order to make the patient acquainted with the navigation and with controls.

Final evaluation: ++++

Critical aspect: Interaction – difficulty in positioning correctly in front of the object where the interaction is foreseen

On the basis of the previous experiences (VREPAR Project) it was decided not to allow patients to manage the controls concerning interaction, as it was noticed that patients too frequently used them incorrectly. When patient focuses an object, the mark on the screen allows the therapists to activate the interaction (doors' and cupboards' opening, objects moving). This procedure is aimed at protecting the patient's VR experience allowing an easier monitoring of some pathological processes (i.e. binge eating). Moreover, by displaying a specific positioning in front of the selected objects, a correct visualisation is possible (i.e. food on the table).

Final evaluation: ++++

Critical aspect: Interaction – Absent in the weighing machine room

Considering the specific characteristics of patients who avoid facing real data concerning their own problem, it is up to the psychotherapist to insert data concerning weight (previously collected) while the patient can read his weight on the display while the index of the scale remains unvaried.

Final evaluation: ++++

Critical aspect: Interaction – absent in the Supermarket scenario

There is a lack of interaction in the environment owing to technical reasons. It would not have been cost-effective to make it possible to have an interaction with food which would require relevant graphic adaptations. The clinical observation criterion was **to discriminate between 'anxiogenous pathways' and 'neutral pathways' according to the typology of the products**:

"...we put an anxiogenous department near to a non anxiogenous department so that if you are anxious you can easily move your attention from one department to the other"

Final evaluation: ++++

Critical aspect: difficulties in movements caused by external devices

The navigation experience in the therapeutic situation is easier than the one experienced in the usability tests, as in the clinical phase any difficulty that may occur in controls is mastered and in some cases anticipated by the therapist. A few critical aspect still remain, which are connected with the use of the headset and with the possibility of physical movements in hampered patients

"...we help them with the mouse. It happens in some environments as, for example, in the supermarket, where you have to follow passages or revolve 180° and anyway the patient doesn't feel belittled or clumsy...we make him feel normal"

During the session the therapist can touch the patient to help him in rotation movements characterising the navigation mode in some environments (BIVRS)

"...so you can touch him, for example, in BIVRS you can help me to find the correct position. It is also an emphatic contact: the patient knows that you are there. Sometimes we (therapists) get up and look towards the headset or to the monitor to help them"

Final evaluation: +++

Critical aspect: Graphic appeal

In contrast with the results emerged in *phase 2*, patients affected by eating disorders give no importance to the graphic attractiveness of the environments and this indicator plays a role of minor importance in the context of the therapeutic process. Patients need a very basic representation of the stimulus to reactivate the non-functional behavioural mechanism. The attention they pay to the vividness, the completeness of the scenario, the realism, the graphic details of the stimulus is inversely proportional to the gravity degree of the pathology. If there are serious difficulties in interpreting the environment, the therapist helps the patient.

"...from a clinical point of view, the realism makes no difference: there is no correlation between effectiveness of the therapy and the realism of the environment"

Final evaluation: ++++

4.3.2 Outline of the semi-structured interview

a) Use of the modules

- Can you tell us which were the patients' selection criteria?
- Which modules did you use? With reference to which pathologies?

- As far as you know, during the clinical trial, there have been deviations with respect to the protocol?
- Was the clinical trial brought to conclusion according to scheduled times and modalities?

b) Expectations with respect to the protocol

- What kind of expectations did the therapeutic protocol give rise to? (as for your experience).
  Notes for the interviewer
  - *Clarify that in this case we mean 'integration of VR in the therapeutic session'*
  - *Stimulate references with specific pathologies and examples*
- Did you think about possible resistances of the patients?
- And what about your colleagues involved in the project?
  Notes for the interviewer
  - *Stimulate considerations about possible personal difficulties / resistances referred to the control of the therapeutic session*
- Which are the main changes with references to the therapeutic protocols you usually use?
- How did you think patients should have faced VR during their therapeutic sessions?

c) Usability

c.1) <u>Local interaction with the artefact</u>

- To evaluate the ergonomic aspects of the modules you used, we considered some indicators. We would like to have your opinion about them. We invite you to think about each of them with reference to the entire system. If you like, you can make specific examples
  The interviewer reads the indicators' list (navigability, contents' completeness, etc.) adding a short description. He makes specific reference to the navigation toolbar only with therapists who used scenarios with command menu appearing on the screen
- The patient is not free to interact with objects, but he's guided by the therapist. How is the navigation controlled in this specific case? Which specific commands are controlled by the patient and which ones are controlled by the therapist?
- In your opinion and with reference to the specific pathology you treat, which are the scenarios you judge more useful (and why) and which ones seem to be less effective? What is your 'satisfaction degree'?
- Let's talk now about patients. How much did they seem involved in the different situations?
- Did they seem sensitive with respect to the graphic appeal? And with respect to other indicators? (*the interviewer reads the list again trying to stimulate a possible confrontation with common users)*
- Which was the patient's reaction when a problem connected to the system occurred (i.e. blocks, slowing down, etc.)?

c.2) Interpretation of the situation

- Taking into account the above mentioned considerations, what was, in your opinion, the VR influence during the therapeutic session on:

- Your job, as therapist?
- The patient's behaviour?

- According to you, what do patients think about the introduction of VR in the therapeutic session?

c.3) Definition of the context

- In your opinion, which was the incidence of the use of VR modules on the progress of the therapy?

d) Towards a culture of use

- How do you think other colleagues sharing your same theoretical approach could use the VR in the context of their therapies?
- Do you think that also colleagues with different theoretical approaches could integrate VR in their therapeutic sessions?
- Which are the main steps for a possible professional training for therapists?

5. Conclusion

On the basis of the evaluations carried out, the usability requirements for a correct and effective use of the VEPSY Eating Disorders modules are fulfilled. It was possible to verify that critical aspects according to the classic usability parameters were faced and effectively solved in the context of the clinical application. On the other hand, some lacks revealed as 'plus' when put in the context of the therapeutic framework as in the case of the graphic appeal, that does not influence effectiveness of the VR therapy at all. The VR scenarios have the function to speed up the access to the personal experience of patients affected by specific psychopathologies and the representation of the stimuli functional to the activation of this process does not need to fulfil requirements connected with the realism of the experience intended as focus on the physical characteristics of VEs: "In this sense, emphasis shifts from quality of image to freedom of movement, from the graphic perfection of the system to the actions of actors in the environment"[25]. Through a correct interaction between the therapist and the patient it is possible to anticipate and avoid any problem of orientation and navigation. The use of devices is simplified and the system is accessible.

From a theoretical point of view it is clearly necessary to draw up new methodological criteria for the evaluation of VR usability minimum requirements when the aim is not the perfect realism of the stimuli represented. Starting from the evaluation of the critical level with reference to usability indicators usually applied, the investigation of the real context of use is the fundamental step that can effectively contribute to the optimisation of the whole designing processes together with the effective integration of existing methods and the improvement of usability evaluation's tools that are still too vague for VR applications

At the end of this paper, we can say that the criteria adopted to analyze VEPSY VEs allowed us to achieve the following:

- Recognize the mediated character of every experience of presence
- Conceive the experience of artefacts' use as immersed in a social context and goal-driven
- Stress the component of ambiguity inherent in everyday situations
- Demonstrate how cultural dimensions affect the effective use of VEs.

The research has thus become a moment of construction - and not of simple application - of new tools for the analysis. One of its main features is that it does not have only a descriptive function, but it also contributes to prescribe solutions to make the VEs systems more efficient, to reach, in every specific case, a sufficiently concrete improvement of the whole interactive process, abandoning an artefact-centered perspective or one restricted to the user–artefact interaction with the main focus on functional characteristics only. And developing this perspective will be our engagement for future usability analysis.

6. Acknowledgments

The present work was carried out in the framework of the European Project VEPSY UPDATED, IST-2000-25323 supported by the European Commission.

References

[1] G. Mantovani. *Ergonomia. Lavoro, sicurezza e nuove tecnologie*, Bologna, Italy, Il Mulino, **2002**.

[2] A.G. Sutcliffe, B. Gault, & J.E. Shin, Heuristic evaluation of virtual reality applications, in: *Interacting with Computers*, **2003**, **in press**, pp. 1-2.

[3] M. Slater, V. Linakis et al., Immersion, presence and performance in virtual environments: An experiment using tri-dimensional chess. Available on line: *www.cs.ucl.ac.uk/staff/m.slater/Papers/Chess/index.html,* **1996**.

[4] C. Zucchermaglio, *Psicologia culturale dei gruppi*, Roma, Italy, Carocci Editore, **2002**, p. 122.

[5] M. Cantamesse, L. Menti, L'usabilità dei siti web: una sfida per le istituzioni universitarie, Vita e Pensiero, Milano, Italy: 5, **2002**.

[6] G. Riva. Web usability revised: a situated approach. *Psychnology Journal*, 1:1. Available on line : http://www.psychnology.org, **2002**.

[7] P.W. Jordan, *An introduction to usability*. Taylor and Francis, **1998**.

[8] A.Pedon, *Metodologia per le scienze del comportamento*, Bologna, Italy, Il Mulino, **1995**.

[9] J.L. Gabbard, K. Swartz K. Richey, D. Hix, *Usability evaluation techniques: A novel method for assessing the usability of an immersive medical visualization VE,* VWSIM '99 Proceedings, San Fransisco, CA, **1999**, pp. 165-170.

[10] J.L. Gabbard, D. Hix, *A taxonomy of usability characteristics in virtual environments: Deliverable to Office of Naval Research, grant no. N00014-96-1-0385*, Blacksburg VA: Dept of Computer Science, Virginia Polytechnic Institute, **1997**.

[11] D.A. Bowman, L.F. Hodges, An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments, *Proceedings: 1997 Symposium on Interactive 3D Graphics*, **1997**.

[12] C. Johnson, On the problems of validating desktop VR, *People and Computers XIII. Proceedings: HCI '98, Sheffield 1-4 September 1998*, Berlin: Springer-Verlag, **1998**, pp. 327-338.

[13] R.S. Kalawsky, VRUSE: A computerised diagnositc tool for usability evaluation of virtual/synthetic environment systems, *Applied Ergonomics*, *30*1, **1999**, pp. 11-25.

[14] D. Hix, J.E. Swan, J. L. Gabbard, M. McGee, J. Durbin, T. King, User-centered design and evaluation of a real-time battlefield visualization virtual environment, in L. Rosenblum, P. Astheimer, D. Teichmann (Eds.) In *Proceedings: IEEE Virtual Reality '99, Houston TX 13-17 March* **1999**, Los Alamitos CA: IEEE Computer Society Press, pp. 96-103.

[15] D.A. Bowman, D.B. Johnson, L.F. Hodges, Testbed evaluation of virtual environment interaction techniques, *Proceedings: ACM Symposium on Virtual Reality Software and Technology, London 20-22 December* **1999**, New York: ACM Press, pp. 26-33.

[16] J.L. Gabbard, K. Swartz, K. Richey, D. Hix, Usability Evaluation Techniques: a Novel Method for Assessing the Usability of an Immersive Medical Visualization VE, in *VWSIM'99 Proceedings*, San Francisco, **1999**, pp.165-170.

[17] C. Zucchermaglio. *Psicologia culturale dei gruppi*. Roma, Italy: Carocci Editore, **2002,** pp. 50-51

[18] J. Nielsen, R.L. Mack, *Usability Inspection Methods*, John Wiley & Sons, New York, NY, **1994**.

[19] L. Gamberini, A. Spagnolli, D. Gasparini, Situated breakdown analysis for the evaluation of a virtual environment, *Psychnology Journal*, 1:1. **2002**. Available on line: http:// www.psychnology.org

[20] L. Gamberini, E. Valentini, Web Usability Today: Theories, Approach and Methods, in: G. Riva, C. Galimberti (Eds) *Towards CyberPsychology. Mind, Cognition and Society in the Internet Age*, Amsterdam, The Netherlands : IOS Press, **2001**, pp. 109-125.
[21] J. Nielsen, R.L. Mack. *Usability Inspection Methods*, John Wiley & Sons, New York, USA, **1994**.
[22] C. Galimberti, G. Riva, *Towards Cyberpsychology. Mind, Cognition and Society in the Internet Age*. Amsterdam, The Netherlands : IOS Press, **2001**.
[23] G. Riva et al. Virtual Reality as Communicative Medium, in : G. Riva, and F. Davide, (Eds) *Communications Through Virtual Technologies*, Amsterdam, The Netherlands : IOS Press, **2001**.
[24] A. G. Sutcliffe, B. Gault and J-E. Shin, Heuristic evaluation of virtual reality applications, *Interacting with Computers*, **2003 in press,** pp. 3-5.
[25] J.L Gabbard, D. Hix, *A taxonomy of usability characteristics in virtual environments: Deliverable to Office of Naval Research, grant no. N00014-96-1-0385*, Blacksburg VA: Dept of Computer Science, Virginia Polytechnic Institute, **1997**.
[26] G. Riva et al., Virtual Reality and telemedicine based experiential cognitive therapy : rationale and clinical protocol, in :G. Riva, and C. Galimberti (Eds) *Towards CyberPsychology. Mind, Cognition and Society in the Internet Age,* Amsterdam, The Netherlands : IOS Press, **2001,** p. 288.
[27] S. Bagnara, *L'ergonomia del software*, Padova, Italy, Clup, **1987**.
[28] L. J. Bannon, From human factors to human actors: the role of psychology and HCI studies in system design, in: Greenbaum and Kyng (Eds). *Design at work: cooperative design of computer systems*. Hillsdale, NJ: Erlbaum, **1991**.
[29] D.A. Bowman, J. Gabbard, D. Hix, A survey of usability evaluation in virtual environments: Classification and comparison of methods. *Presence: Teleoperators and Environments*, *11*4, **2002**, pp. 404-424.
[30] D.A Bowman, L. F. Hodges, Of interacting techniques for immersive virtual environments. *Journal of Visual Languages and Computing*, 10, 37-53, **1999**.
[31] J. Bruner, *Acts of meaning*, Cambridge, MA: Harvard University Press, **1990**.
[32] J. M. Carroll, R. L. Mack, S. P. Robertson, and M. B. Rosson, Binding objects to scenarios of use. *International Journal of Human-Computer Studies*, 41, **1994**, pp. 243-276.
[33] C. Galimberti, G. Riva, *La comunicazione virtuale. Dal computer alle reti telematiche: nuove forme di interazione sociale,* Milano, Italy, Guerini e Associati, **1997**.
[34] N.V. Hammond, P.J. Bernard, J. Morton, and J.B.Long, Characterizing user performance in command-driven dialogue, *Behaviour and Information Technology*, 6, 2, **1987,** pp.159-205.
[35] P.W. Jordan, *An introduction to usability*, Taylor and Francis, **1998**.
[36] K. Kaur, A.G. Sutcliffe, N.A.M. Maiden, Towards a better understanding of usability problems with virtual environments, in: A. Sasse, C. Johnson (Eds.), *Proceedings of INTERACT 99: IFIP TC.13 Conference on Human Computer Interaction*, Amsterdam: IFIP/IOS Press, **1999**, pp. 527-535.
[37] S. Krug, *Don't make me think! A common sense approach to web usability*, Indianapolis: Macmillan, **2000**.
[38] G. Mantovani, *L'interazione uomo-computer*, Bologna, Italy, Il Mulino, **1995a**.
[39] G. Mantovani, *Comunicazione e identità. Dalle situazioni quotidiane agli ambienti virtuali,* Bologna, Italy, Il Mulino, **1995b**.
[40] D. H. McBurney, *Metodi della ricerca in psicologia*, Bologna, Italy, Il Mulino, **1996**.
[41] M.B.Miles, A.M. Huberman, *Qualitative data analysis (second edition),* London, UK: Sage, **1994**.
[42] B. Nardi, *Context and consciousness. Activity Theory and Human-Computer Interaction*, Cambridge, UK: The MIT Press, **1996**.
[43] J. Nielsen, Usability engineering at a discount, in: G. Salvendy, M. J. Smith, (Eds.) *Designing and using human computer interfaces and knowledge based systems*, Ansterdam, NL: Elsevier, **1989**.
[44] J. Nielsen, *Usability Engineering*, San Diego, CA: Academic Press, **1993**.
[45] J. Nielsen, *Designing Web Usability*, Basingstoke, England**:** Macmillan computer publishing, **2000**.
[46] I. Poupyrev, T. Ichikawa, Manipulating objects in virtual worlds: Categorization and empirical evaluation of interaction techniques, *Journal of Visual Languages and Computing*, *10*1, **1999**, pp. 19-35.
[47] G. Riva, Web usability revised: a situated approach, *Psychnology Journal*, 1:1, **2002**. Available on line: http://www.psychnology.org
[48] M.J. Schuemie, *Human–Computer Interaction and Presence in Virtual Reality Exposure Therapy*, PhD dissertation, Technische Universiteit Delft, Delft, **2003**.
[49] L. Suchman. *Plans and situated actions*, Cambridge, UK: Cambridge University Press, **1987**.
[50] F.Sudweeks, S.J. Simoff, Complementary explorative data analysis: the reconciliation of quantitative and qualitative principles, in: S. Jones, (Ed.) *Doing Internet research: critical issues and methods for examining the Net,* Thousand Oaks, CA: Sage Publications, **1999**, 29-56.
[51] A.G. Sutcliffe, K. Kaur, Modelling interaction for virtual reality systems. *Poster paper: INTERACT'97, Sydney*, Amsterdam: IOS Press, **1997**.

[52] A.G. Sutcliffe, K.D. Kaur, Evaluating the usability of virtual reality user interfaces, *Behaviour and Information Technology*, *19*6, **2000**, pp. 415-426.
[53] A.G.Sutcliffe, *Multimedia and virtual reality: Designing multisensory user interfaces*, Mahwah NJ: Lawrence Erlbaum Associates, **2003**.
[54] M.Visciola, *Usabilità dei siti web*, Milano, Italy, Apogeo, **2000**.
[55] C. Wharton, J. Reiman, C. Lewis, P. Polson, The cognitive walkthrough method: A practitioners guide, in: J. Nielsen, R. L. Mack (Eds), *Usability inspection methods*, New York: Wiley, **1994**, pp. 105-140.