Narcis Parés*
Roc Parés
Experimentation on Interactive Communication, Audiovisual Institute
Universitat Pompeu Fabra
Barcelona, Spain

Towards a Model for a Virtual Reality Experience: The Virtual Subjectiveness

Abstract

After analyzing how VR experiences are modeled within human computer interaction (CHI) we have found there is a deep theoretical gap. Similarly to how the scientific community has defined CHI models for multimedia applications, it would be very important to have such models for VR—obviously the standpoint cannot be the same because multimedia and VR applications differ in essence. Indeed, there is no formal model to unify the way in which scientists and designers of VR applications define their experiences. More specifically, apart from the isolated initial scheme defined by S.R. Ellis (Ellis, 1991, *Computing Systems in Engineering, 2*(4) 321–347; Ellis, 1993, *Pictorial Communication in Virtual and Real Environment,* 3–11), and a low level model defined by Latta and Oberg (Latta & Oberg, 1994, *IEEE Computer Graphics & Applications, 14,* 23–29), there is no model to fully describe the relationship with which the user will be experiencing the VR application.

In this paper we shall explain where exactly we think this gap is found, which elements and concepts are involved in the definition of a model of experience, and finally propose a definition of a model that we think, eventually, will fill this gap.

1 Introduction

Scientists have been trying to find and define the full potential of VR for the last thirty years. Many have restricted their search to the field of simulation. Others have observed that VR is more than a simulation technology and have widened their view to study the full spectrum of fields (e.g., Biocca, Kim, & Levy, 1995; Parés & Parés, 1993, 1995; Steuer, 1992; Stone, 1993). The lack of a formal, timeless, lasting definition of VR, together with the strong emphasis on technology, have, in our opinion, hindered the definition of a model that expresses fully and coherently the relationship between a user and a VR experience. Such a model would help and guide designers, scientists, and developers involved in VR to use this technology/medium in a justified and rich way. This would hopefully minimize applications where VR is used with a sensationalist and/or poor approach. For instance, it could help leisure creators in understanding how to use VR without falling in the temptation of using it as a substitute for, say, cinema or theater. It could also help scientists in

Presence, *Vol. 15, No. 5, October 2006,* 524–538

*Correspondence to npares@iua.upf.edu

designing the type of experience they really want the user to have, for her to correctly understand a concept, get a task done, or get trained.

In this paper we will take a view of VR not as a mere technology, but rather as a medium—a communication medium in the sense of it being able to convey meaning, to transfer knowledge, and to generate experience. We appreciate some efforts done in this direction (e.g., Steuer, 1992), but we want to differentiate our position. Specifically, we do not see it from the mass media approach as Biocca et al. do (Biocca & Levy, 1995; Biocca et al., 1995), nor do we see VR as a telepresence-based communication medium as Steuer (1992) does. Therefore, our understanding of VR will be that of an interactive communication medium that is generated in real time. The intention behind approaching the analysis of VR as a medium is to be able to explore its potential distinctive properties independent of the specific technology of the moment. This, on the one hand, should avoid falling in past situations such as when VR was thought to include only those applications that used an HMD and a data glove; definitions that have later been found obsolete and have had to be reformulated to include CAVEs and many other technologies. On the other hand, if any true contribution is to be made by digital media, it must come from exploiting new and intrinsic qualities and not from imitating or substituting what can already be done in other media (Brown, 1994; Parés & Parés, 1993, 1995). Also we want to define the model of VR without being constrained by current applications in order to leave all future possibilities intact. In this sense we will not restrict our approach to simulation uses of VR.

2 Where is the Gap?

In the history of VR, S. R. Ellis (1991, 1993) defined what we could consider the strongest attempt to formalize a structural model of a VR application and the elements that participate in this structure. According to Ellis, an environment is composed of three parts: content, geometry, and dynamics. Many software libraries and tools have used this model as a basis for their user-application paradigm and, although some have slight variations, the elementary concepts have not changed significantly (e.g., WorldToolKit from Sense8, Inc., 1999; DIVE from the Swedish Institute of Computer Science, Carlsson & Hagsand, 1993, etc.). However, this approach models only the structure of the virtual environment and does not model the "experience" of the user; that is, it does not fully describe how the VR is mediated to the user. Indeed, Ellis defines content as a set of objects and actors that are themselves defined by state vectors which identify all their properties (physical, structural, etc.). He then distinguishes a particular actor that he calls the *self*. This self is "a distinct actor in the environment which provides a point of view from which the environment may be constructed" (Ellis, 1991, p. 322). In this definition we see that there is an initial attempt to define a structure that describes the user experience. However, this structure is not complete in the sense that he defines it as an internal element of the environment that is not specifically linked to the exterior (the user). We shall detail further these aspects below, in our model.

On the other hand, Latta and Oberg (1994) defined their "Conceptual VR Model" in such a way that although they do emphasize a user-centered approach, they stay at a very low abstraction level of the perception of the user and therefore the experience is only defined at a sensorial and effector level. In other words, the experience is analyzed from the point of view of the perception of digital stimuli by the user, and how the motor system of the user may influence the management of the application. Figure 1 summarizes their schemes from the human view and the technical view of the VR system.

In this scheme we can see how they focus on the interface at a very low description level and remain very much centered on the physical interfaces.

These were very interesting first approaches, however, ten years and many application areas, new technologies, and new definitions of VR have gone by since then. The gap we have found is therefore at a high level of description. We think it is important to obtain a model of how the user experience is mediated in a VR application; what can make the experience different, how can the

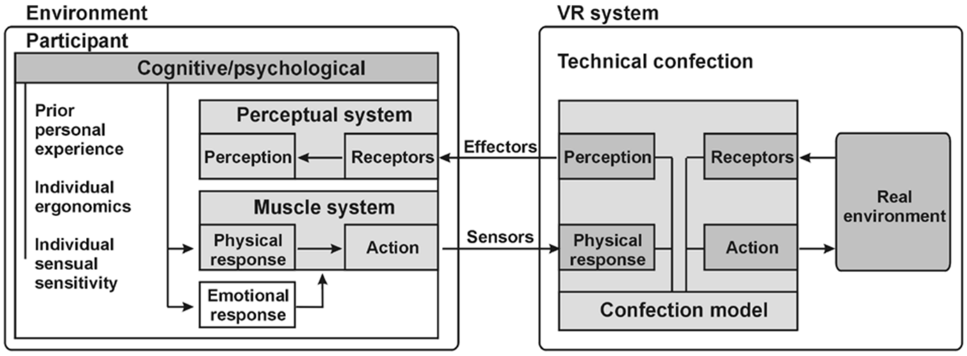

Figure 1. *Diagram summarizing the human view and technical view of a VR system by Latta and Oberg (1994).*

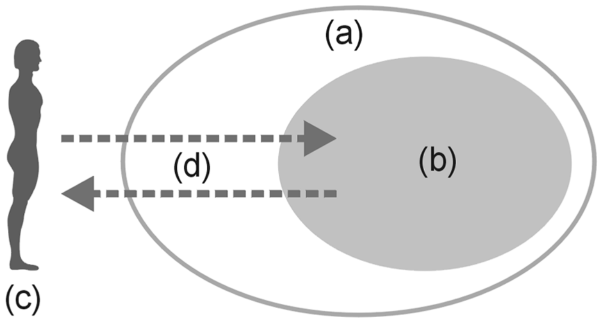

Figure 2. *Simple CHI diagram of (a) VR system, (b) virtual environment, (c) user, and (d) interactive communication.*

user understand it and how he can influence the states and actions of the application.

So, how can we typify this gap? Figure 2 shows a very simple diagram of the CHI structure that relates a VR system (a) and a user (c). The user is related to the system by a two-way communication; that is, an interactive communication (d). It is therefore important to stress that the user is not related to the VE (b) by a mere action/reaction type of interaction. This is a key change in focus because by putting the accent on communication, as opposed to technology, we are explicitly referring to the exchange of information through a medium. This means that some sort of filtering action is giving a specific view of the communication experience to the user: the mediation. Hence, depending on how this mediation is defined, the user will *understand* the whole communication experience in one way or another.

To provide a simple metaphorical example, Figure 3 shows a VR system (a) that holds the definition of a specific VE (b). As designers, we could define a filter within the system that would force the user to experience the environment in a particular way. In our example, we can represent this filter as a pair of virtual eyeglasses (c) that present the VE (b) to the user (d) with a vertical pattern of lines. This user would then have the impression that the VE has indeed this peculiar visual presentation (e). On the other hand, we could define another filter (f) that would force another user (g) to think that VE (b) has a peculiar horizontal pattern associated to it (h). It must be stressed that the VR system (a) and, more importantly, the VE (b) are the same in both cases. This means that although the kernel definition of the VE may be unique, different users may have different experiences.

This example obviously oversimplifies the true situation. We need not specifically design a pair of virtual eyeglasses to give the user a filtered understanding of the experience. We could rather define a specific viewing range or a particular navigation speed and we would also be affecting the understanding of the VE by the user. Notably, notions such as transparent interface or nonmediated experience have imbued many VR designers, developers, and even toolkit paradigms with the false idea that they can actually define neutral VR sys-

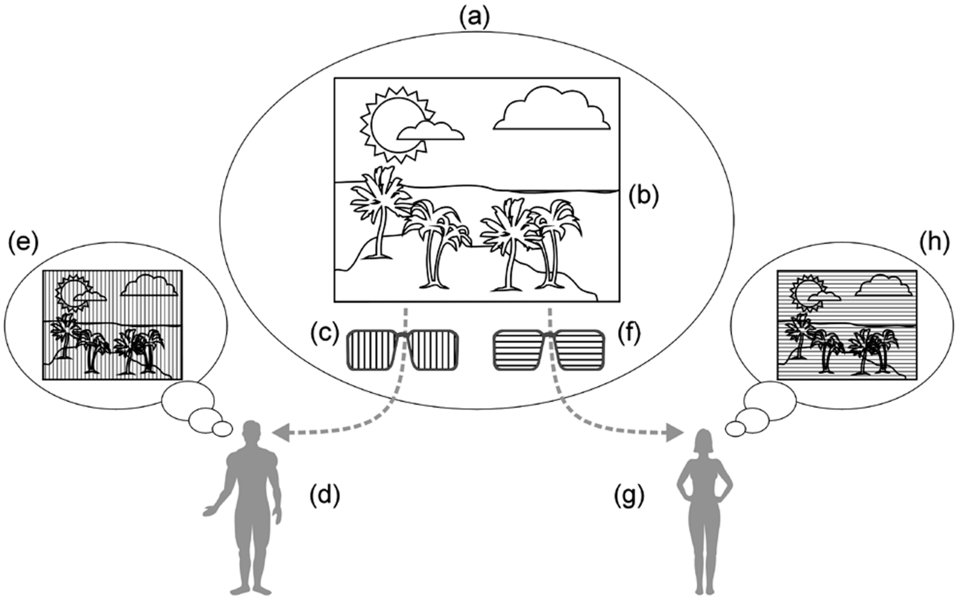

Figure 3. *Simple metaphorical example of a filter that modifies user understanding in the communication experience provided by the VR system (see text for details).*

tems. By neutral we mean that they believe these VR systems are merely simulating a real experience. Hence, on having in mind the goal of eventually achieving experiences that are indistinguishable from reality, many believe these neutral systems and transparent interfaces will not mediate the experience to the user. From a communication standpoint this has been proven false. It is important to remember that "mediation will not occur in communication because a medium exists in-between, be it technological (e.g., TV) or natural (e.g., the air), but rather because through a technological medium, a mediator constructs a specific discourse" (Rodrigo, 1995). This not only confirms the fact that design decisions affect mediation, but also emphasizes the important role that designers and developers have in experience design. In any case, a VR system that would successfully provide experiences indistinguishable from reality would not only need to create realistic VEs but also would need to mimic the human perception and reception of it in what could be called a human-like virtual subjectiveness (VS).

What we are trying to model in this paper is this mediation. This is the actual gap we have found and which we will call the VS because it gives the notion of the user having a subjective view of the experience due to the mediating action described. Therefore, if we can formalize this mediation and we can understand how it works, we will better design and develop experiences for users.

3 The VR Experience

To begin the definition of our model, we would like to propose a differentiation between the terms virtual environment and virtual reality. These terms have been historically used as synonyms; however, to our understanding, having so many terms that are used with little, if any, distinction in our field (VE, VR, telepresence, cyberspace, etc.) only impoverishes the field and causes confusion. Therefore, we propose a unique meaning for each term that we believe will enhance comprehension and concept clarification in our scientific community.

To aid us in this differentiation, we will use an analogy with two concepts from the field of simulation. According to Whicker and Sigelman (1991) a *model* of a simulation is a representation of the structure to be simulated; that is, a *static* definition that establishes structures, parameters, and functions or algorithms. On the other hand, the *simulation* itself is a representation of such structure *in action;* that is, when the model of a simulation is made to evolve over time, starting from an initial state, feeding it with input information and obtaining an output that is (hopefully) the desired result.

The analogy we propose is that a virtual environment (VE) be equivalent to the model of a simulation. Consequently, we propose that it only refers to *static* structures, that is, a VE would include Ellis' structural description of content (object definition), geometry (numerical database) and dynamics (static set of rules of the environment). On the other hand, we propose that virtual reality (VR) be the structures of a VE *put in action*. In other words, VR would be equivalent to *simulation* in the sense that it would refer to the case in which the VE is made to evolve over time. Therefore, VR is the real time experience a user can have of a VE. In this sense, a VE could be used for a real time experience (interactive or not) or it could be used for a non-real time, offline rendered CG animation, or for a CG single image rendering (e.g., for a poster). Hence, a VE is not, by itself, associated with the user (in that it does not provide or generate anything) until it is put in action and interfaced with the user, so that it evolves over a period of time, during which the user perceives it and interacts with it.

We would like to stress here that with this analogy we are by no means trying to restrict VR to the area of simulation or to simulation applications. On the contrary, this paper will hopefully widen the definition of VR to encompass all those applications of VR that lie outside the area of simulation and that are often left out of most existing definitions—for example, most of the work developed by Myron Krueger and his Video Place system (Krueger, 1985), which is widely acknowledged as being part of the realm of VR but is systematically left out by almost all definitions of VR.

Going back to the definition of a VE given above, we have seen that it fits exactly into what Ellis defined as content, geometry, and dynamics. However, we then encounter the key questions that this paper will try to answer: How can a user have an experience of this VE? How can the user understand and interact with the VE during a period of time? How is the VR experience generated and, more importantly, mediated to the user?

Of course we could take the approach of Latta and Oberg of studying how displays generate stimuli for the user, how sensors capture user actions, and how the perceptual and muscle systems of the user react. However, as we have seen, this approach is too focused on the physical interfaces that link the user and the VR application; that is, they are too focused on technology. We believe there is a need for a higher order conceptual model that relates more to the *semantics* of the experience rather than to the perception and facilitation of the experience.

Another point we would like to clarify and stress in this paper is that our focus is on stimuli generated by a computer system. Therefore, we would like to differentiate our view of VR from any concept related to telepresence—as opposed to Latta and Oberg (1994) or Steuer (1992). We understand telepresence in its original sense defined by Marvin Minsky (1980). The rationale behind this is that on designing a telepresence application, one may have very low control over the mediation stated in the previous section. This is because it is an application that translates a physical world, the one on which the user will operate remotely, onto another physical world, that in which the user is found. In this context, we would like to define VR as interaction with real time generated digital stimuli. Although it may sound like a very open definition, we see three advantages in it. The first is that it does not restrict VR to 3D experiences and therefore, as stated previously, includes important 2D VR work such as that done by Krueger, the Vivid Group (Vincent, 1993), and others. The second advantage is that it puts the accent on digital stimuli and therefore does not tie the definition to a specific VR technology. Lastly, it reinforces the idea of real time generation for the experience, which is different from real time interaction. The idea of having to generate responses in real time (on the fly) is unlike any other medium. This is an important property as we describe in the following section.

One final point related to the VR experience. The virtual subjectiveness we are trying to model has the goal to help us understand how to correctly define interesting and powerful experiences independent of any goal related to sense of presence. We believe this is important for a number of reasons. On the one hand, the sense of presence is not always desirable (Ellis, 1996, p. 248) and no correlation has been yet proven to exist between task performance and sense of presence (e.g., Slater & Wilbur, 1997). Also, presence is a term intimately linked to the fidelity of stimuli (e.g., Sheridan, 1996) and this suggests that it may only be applied to simulation applications (explained and supported by Ellis, 1996, p. 248). Therefore, it seems as if those applications that lie outside the realm of simulation cannot yield a sense of presence because there cannot be fidelity towards an imaginary world (in such applications it would probably be more appropriate to talk about a sense of agency[1]). Therefore, if our virtual subjectiveness model were linked to sense of presence we would not be providing a general purpose model.

4 CHI for Multimedia and VR

We have briefly mentioned that CHI models must be different in VR to those for multimedia (MM) applications because they are different in essence. Therefore, before we go into our model, we would like to clarify how we think MM and VR differ.

MM is based on the integration of digitized media: images, sounds, videos, and so forth. This integration defines a situation in which all the options and data (or media) are prerecorded; that is, there is no real time generation of the material. On the other hand, VR is based on underlying mathematical models (not necessarily models of our physical reality) in the computer that require real time management to generate the desired final stimuli: images, sounds, and so forth (see Figure 4).

This is important, more than at a perceptual level, at an interaction level. In other words, at a perceptual

[1]The sense acquired when the user is conscious of being able to exert control over his surrounding environment.

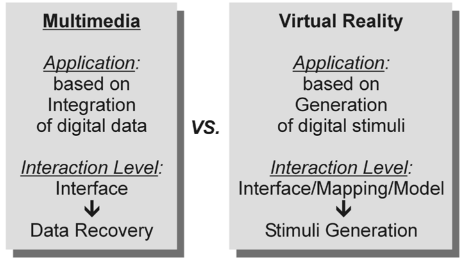

Figure 4. *Essential differences between multimedia and virtual reality applications from a CHI standpoint.*

level, the user can barely differentiate whether the application is MM or VR because we are only concentrating on how those stimuli are interfaced with our sensorimotor systems. Hence, the difference between MM and VR is important at an interaction level because a user of MM applications is confronted with a situation that can be generally described as *data recovery,* whereas the VR user is confronted with a situation of *real time stimuli generation* that is guided by three main components: interfaces, mappings, and model. In other words, in MM applications the user searches for information and recovers it from different places within the application where the information is structured and predefined. In VR applications, the user experiences situations that are generated by her interaction and which allow her to explore, manipulate, or contribute to the experience. Of course nowadays many hybrid applications may be found defining a continuum from MM to VR. However, for the sake of clarity we feel it is important to define a formal boundary between the two. Let us now define how this interaction occurs in VR.

5 The Interface

Before going into the kernel of our proposal, we would like to just briefly clarify one more point. This is the term and concept of *interface.* It is also a somewhat confusing term because each researcher uses it in slightly different ways and therefore, we would like to state what we understand by it. It is very common to find people referring to the interface of a system as mainly the physical devices that are related to input; for example, "the interface of the application was a 3D mouse and an electromagnetic position and orientation sensor." This is one of the main problems we see with Latta and Oberg's (1994) model and proposal. Of course sometimes this is only an implicit way of speaking, but this tends to make people stop thinking of the interface as a two-way system (of course, output devices are also interfaces), and, more importantly, it very often causes people to forget about the rest of the interface system: the *logical* (or software) interface and the *mappings.*

For example, we believe it is important to always keep in mind that when we interact with a PC nowadays in a windows environment (a WIMP system; *Windows, Icons, Menus, and Pointing device*), we have the mouse as a physical interface, the cursor as the logical interface, a mapping between the two, and a screen to visualize the results. It may sound obvious, but it must not be forgotten that without the cursor, the mouse is useless. Also, without a proper mapping between the physical and logical units, the functionality of the interface may be useless (either too sluggish or too fast to control).

Therefore, this global idea of the interface must be seen as one of the essential parts of the entity that will allow us to define how the mediation between the user and the experience occurs. Bowman, Kruijff, LaViola, and Poupyrev (2001) present useful guidelines for successful 3D user interface design. However, on the one hand, we do not want to restrict our model to 3D applications. On the other hand, as they state, none of the techniques may be identified as the "best" for all situations because they are task- and environment-dependent. This is the reason why it is important to understand and describe the mediation of the experience to the user through a model such as the virtual subjectiveness.

6 Virtual Subjectiveness: Much More than an Avatar

Ellis gives a first hint of the relation "user-experience" on defining the *self* within his definition of

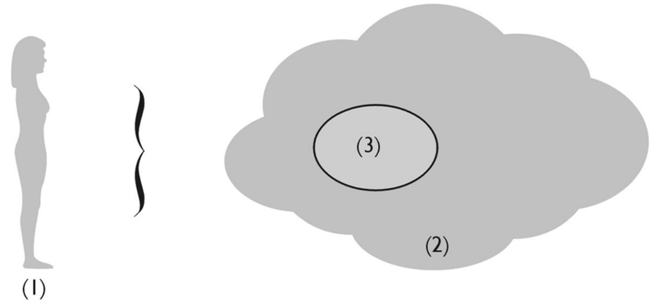

Figure 5. *The user (1), the virtual environment (2) and the self (3), the logical interface.*

content. Nevertheless, it is still an element that is clearly inside his definition of VE.

6.1 Interfaces and Mappings

Let us sketch Ellis' elements in a first model that we will gradually complete. Figure 5 shows a user confronting a VE. In the VE we have the *logical interface,* which under Ellis' nomenclature would be the *self*. The logical interface is, according to Ellis, the definition of the "point of view from which the environment may be constructed" (Ellis, 1991, p. 322). Understanding "point of view" in the widest possible sense, this is for us the first part of the definition of the mediation of the experience between the user and the VE; the logical interface may define the viewing direction, field of view, the type of projection, whether it has stereo view, the hearing capacities, force feedback properties, and so on, and it may define its own appearance or representation (as in the case of an avatar). However, there is no explicit link to the user, or to the actual VE.

Let us incorporate the *physical interfaces* (sensors and displays) to this model and the *mappings* that link them to the logical interface (Figure 6). Now we are not only stating how the VE may be constructed from a point of view, but also we are explaining how this construction is linked to the user's perception (senses) and how the user's motor system is linked to the control of this point of view. This is close to what Latta and Oberg define, although here we would like to stress again the fact that when we say "how the user perceives" the environment,

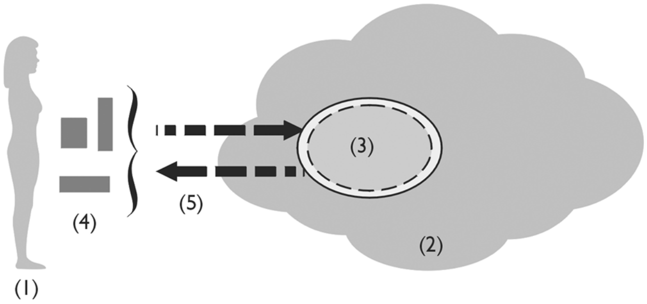

Figure 6. *The user (1), linked to the logical interface (3) within the virtual environment (2) through the physical interfaces (4) according to the mappings (5).*

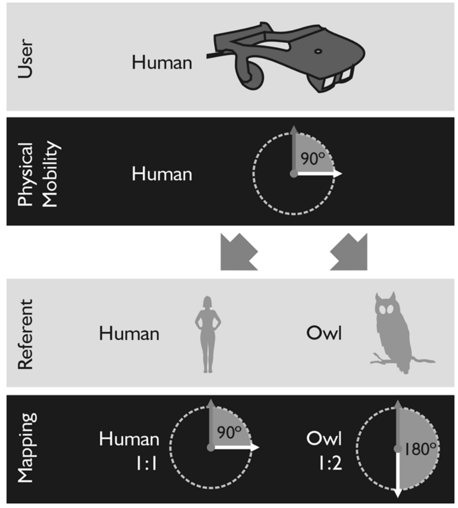

Figure 7. *Two different user to viewpoint mappings for a single VE define different VR experiences.*

we are not referring to the physio-psychological processes, but rather to the cognitive semantic processes.

For example, let us suppose an application where a user is confronted with a VE that defines a forest. Let us suppose that for the physical interface related to visual perception and point of view control, the user is provided with an HMD and a magnetic orientation sensor. Finally let us suppose a reasonable 1:1 mapping between the orientation data of the sensor and the orientation of the point of view of the logical interface; that is, the user turns her head 90° to the right and sees what is virtually 90° to the right of the logical interface. Now, this same forest would be understood by the user in a very different manner if the mapping between the magnetic sensor and the point of view were altered to a 1:2 mapping; that is, the user turns her head 90° to the right and sees what is virtually 180° to the right of the logical interface. This perception, for example, could allow a child to understand how an owl perceives its surrounding world. The owl has the physical capability to turn its head 180°, but the child does not. Of course the user could detect kinesthetic incongruence; however we can see how the definition of a nonstandard mapping permits the possibility of experiencing a single VE in two extremely different ways through a VR experience (Figure 7). It is not a property of the VE—the forest remains the same in both cases. It is rather a property of how the experience is put in action.

Let us analyze another example. Imagine a user on a treadmill (physical interface) linked to the point of view (logical interface). If we define a mapping where, for every physically walked meter the point of view moves 10 units in the virtual environment, we are not giving any clues to what this really signifies to the user and we may be making this definition arbitrarily. If we know the VE is 100 units in any direction, then we know the user needs to walk 10 meters to reach the opposite end of the VE. Suppose this makes any task in the experience very slow and cumbersome because it takes too long for the user (too many steps) to move from one place to the other. Now, apart from the idea of efficiency, cognitively, the user might think she is in a very large environment (Figure 8a). Hence the mapping not only affects her performance and her perception of speed, but it also affects her perception of scale. Let us now modify the mapping such that now, for every 10 centimeters physically walked the point of view moves 10 units. The user might find that not only are the tasks to be undertaken within the experience now too difficult to control because of the huge speed, but also she might feel the environment has suddenly shrunken because she now reaches all edges of the world with no effort (Figure 8b).

This relation defined by the mappings between the physical interfaces—directly linked to the user—and the

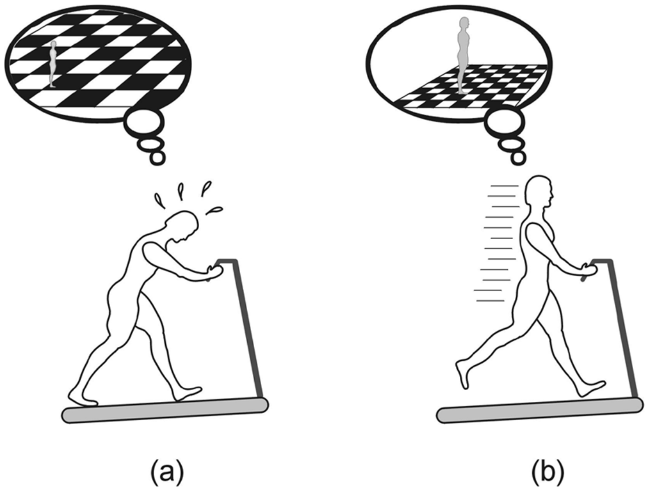

Figure 8. *Different mappings make the user understand the VE in different ways (see text for details).*

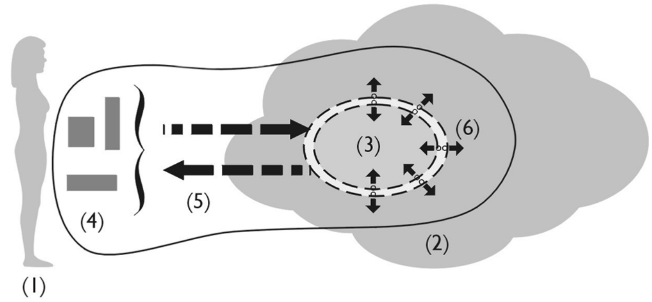

Figure 9. *The virtual subjectiveness emerges from a well-coordinated design, scripting and programming of the physical interfaces (4), the logical interface (3), the mappings, (5) and the behaviors (6), that mediate as a single entity the experience of the virtual environment (2) to the user (1).*

logical interface—indirectly linked to the user—forming a unity of mediation, is extremely rich to work on to define the way we would like the user to perceive the experience. This is why we think this unity must be clearly identified and separated from the definition of the VE in order to help define the experience of the user.

6.2 Behaviors

Now that we have an explicit linkage of the user with the logical interface and hence have a first level of experience, we need a complete linkage of the logical interface with the VE. This comes through the definition of the behaviors that may be associated with the logical interface with respect to the actors and other objects and elements of the VE (Figure 9). These behaviors define explicitly how the user is allowed to interact with the experience through the logical interface. In other words, how the user may affect the VE.

For example, the experience could be an architectural fly-through; that is, it would be an explorative experience where the logical interface, the point of view, has no particular behavior to let the user interact with the VE (except for the real time construction of the images from a point of view). We could start defining behaviors for this logical interface at this point, such as fixing the height of eyesight to a certain distance above the floor and generating collisions with all encountered objects. The experience is still explorative, however, the perception the user has is completely different; she cannot cross walls anymore nor can she have a bird's eye view of the VE. Now let us define the ability for this logical interface to have a manipulation element and be able to move objects in the VE. Again, the perception of the user changes radically even though the VE still has the same definition.

6.3 The Virtual Subjectiveness

At last we come to the final link we needed to define a whole user experience; that is, what we call *the virtual subjectiveness* (VS). Specifically, the VS is the mediation element composed of the physical interfaces, the mappings, the logical interface, and the behaviors (Figure 9). The VS not only mediates the user's experience at a cognitive level, defining for her a specific understanding of the VE, but also defines her complete *unfolding* within the experience. In other words, it not only defines how the user might understand the environment in which she is having the experience, but also what potentiality she has and how she can apply it in her activity and reactions within the experience. Hence, the user does not directly interact with a VE to obtain an experience; rather, the experience of the user is actually mediated by an intermediary entity that we call the VS.

This is the model we propose: a model where the accent is put on the user's relation with the VE—and not on the technological interfaces, nor on the properties of the elements within the VE, nor on the physiological perception of the user. It is a relation based on the semantic information that is transmitted to the user through the VS in the interactive communication process with the VE.

7 The Keystone of VR Experience Development

As we see it, the proposed model represents more than a theoretical advance. It provides a solid framework from which to design and analyze VR experiences. This framework may be summarized from the philosophy described above as a three-layered analysis of design requirements. Although we are still working on this top-down design scheme we present it in Figure 10 as a preliminary reference.

In the case of simulation applications it helps in coherently defining all the properties that the user interaction must have. For example, if the application must define a virtual wind tunnel, our model helps in fully defining the role of the user within such a scientific application; that is, the VS. Hence, it makes explicit the fact that the user might only need be an infinitely small point in space, that defines a point of view (direction and range), and a virtual smoke thread generator, another infinitely small point in space. It might sound obvious, however we often see scientific applications where the actual VS could be extremely simple in its definition, and nonetheless their authors insist on defining anthropomorphic representations and functionalities that limit the actual use of the application. By making it explicit, through our model, the scientist becomes aware of potentiality and limitations that might appear in a specific task that initially were not foreseen. Of course the user always remains a human and is hence limited by its physical and sensorial constraints; however, the adequate definition of the VS, that is, the adequate design

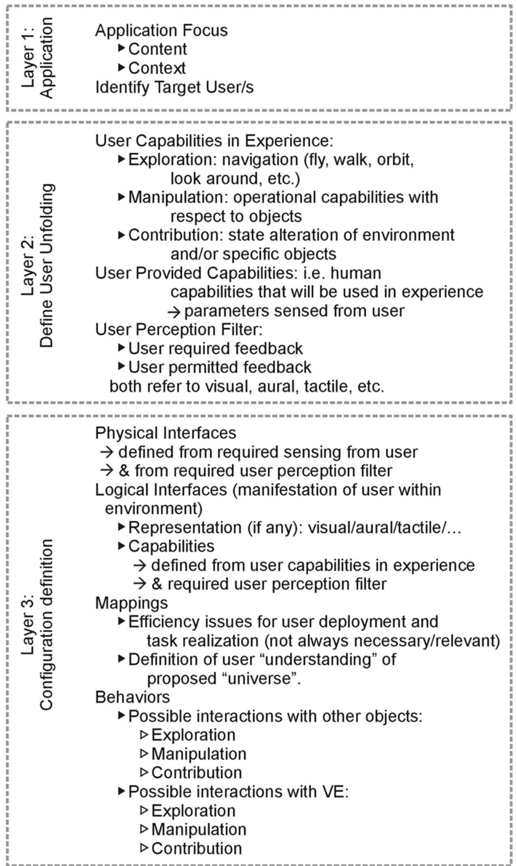

Figure 10. *Preliminary framework definition for top-down VS design.*

of the mediation of the experience, may completely transform the experience.

In the case of leisure or art applications, the VS defines what could be considered as the only, or at least the most powerful, design tool that the “creator” might have during the design process. Let us contrast this with the creative possibilities of a cinema director who has the control of the frame to narrate, convey meaning, describe situations like mystery, and so forth. The director also has control over time through cuts and editing. On the other hand, in VR an experience is, by definition, under the control of the user, because it is interactive and, very importantly, because the generation is in real time; that is, the user chooses what to see, where to go, and when to do so (unless the user is so limited and guided within the application, in which case the experience ceases to be a VR experience or even interactive).

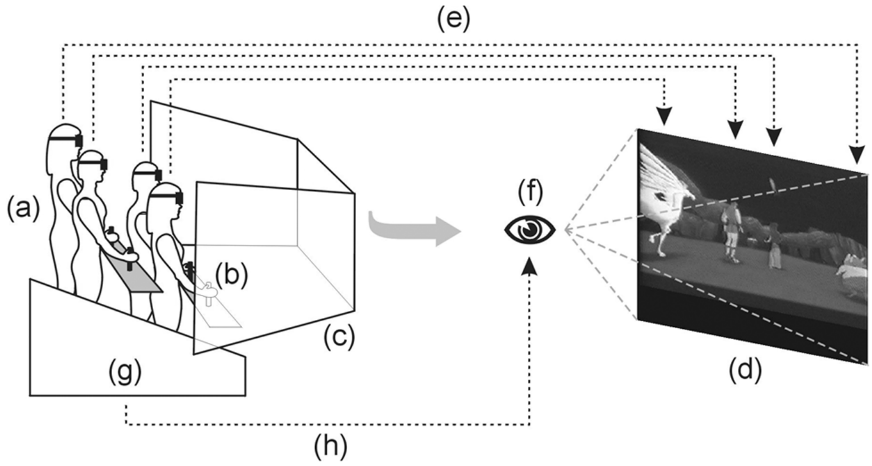

Figure 11. *Schematic diagram of the "Hercules in the Underworld" attraction (see text for details).*

This apparently leaves the VR creator void of creative possibilities. However, the creator still has control over how the user is deployed within the VR application, that is, how the user will unfold its participation, and this is done by designing the VS.

Many researchers have expressed the need to find those elements that relate the user and the experience to better understand what influences the perception, understanding, engagement, and so forth, of the user. For example, we believe this model comes to show that the "form," as defined by Slater (2003) and by Waterworth and Waterworth (2003), in which the stimuli are presented to the user, is of much less importance than the way these stimuli are related to the semantics behind the mediation defined by the VS; mainly because the same VE, with the same presentation form (i.e., sensor/display hardware configuration) may yield a completely different experience to the user by changing the properties of the VS.

8 Analysis of Two Applications

Let us now use our model to analyze two simple example applications to show the importance of a correctly designed VS. The two chosen applications are both from the leisure/entertainment area, specifically, both have been or are VR attractions of the DisneyQuest indoor amusement center at Disney World in Orlando, FL (1998). The first application will serve as an example of a faulty design of the VS, whereas the second will show a correct approach. It is a very interesting and useful situation to be able to have two real applications that are so close to one another because their analysis can then be very clearly and closely compared.

8.1 "Hercules in the Underworld" Attraction

This interactive VR attraction is based on the Disney feature film *Hercules* (Clements, 1997). The idea is that the users embody four characters of the film (Hercules, Phil, Meg, and Pegasus) and fight the evil Hades by gathering thunderbolts from Zeus. Figure 11 shows a schematic diagram of the attraction. In this attraction, up to four users can play (Figure 11a) by each of them interacting through a joystick (Figure 11b) and each wearing LCD shutter glasses to see the stereo images on three screens in front of them (Figure 11c).

Because each user embodies a different character, there is a mapping (Figure 11e) set from each physical interface (the joystick) to each logical interface (the character). The users see their characters in a third person view interaction scheme (Figure 11d). Hence, the user may explore the surrounding environment by moving the character around.

The issue in this setup is that the user in fact, when moving the character, finds himself limited in movement to a certain area around an imaginary central point of the group of characters. Apparently, many users did not understand this limitation and hence, many pushed the joystick strongly to try to move the character further away (causing robustness problems in the physical interface; Mine, 2002).

Let us analyze the situation. On starting the experience, users get the impression that it is indeed a four multiuser game because they find four distinct joysticks and four characters on the screen; that is, users make a reasonable conceptual link between physical input interfaces and logical interfaces. When they make small movements of the character through the joystick, this scheme is reinforced, because of this 1:1 relation, and each user begins to want to explore more of the environment, possibly each in a different direction; that is, a reasonable user understanding of logical interface behavior. However, the mental model (Norman, 1988) that the user applies is not correct because the third person view that the four users have of their character is in fact a common view of "the group" of characters; schematically shown in Figure 11f. What actually happens is that the four users move around the environment as a group and not as individuals. This group cannot divide itself because there is a single set of screens for the four users (there is a mismatch between the apparent amount of logical interfaces and the unique physical output interface). Therefore, when the game leads the users to move in one direction because of the occurring action, they have to all move together in that direction. In other words, the logical interface is actually divided in two from the vision of each user, namely: a unique representation of each character with a limited behavior and a group point of view that holds the basic navigation potential. Between the group of users (the whole set of joysticks; the physical interface; Figure 11g) and the point of view from which the images are generated (the actual logical interface that gives user reference; Figure 11g) a mapping (Figure 11h) is defined that actually reflects the activity VS. This is why the individual characters may have only a limited range of action around this invisible and abstract idea of the group.

In terms of the VS we find it is not a unique entity that clearly defines the experience. The set of interfaces, mappings, and behaviors belong to two clashing definitions or strategies, and hence confuse the user because they are not compatible. One gives the user the impression of being an individual entity that can freely explore and interact with the experience within a multiuser application. The other gives the sense of being a group of users moving together within an environment that may only be explored by the collaboration of the entire group (a group application).

8.2 "Pirates of the Caribbean: Battle for Buccaneer Gold" Attraction

This interactive VR attraction is based on the famous Disney classical attraction "Pirates of the Caribbean" (1967). Here, the idea is that the users become pirate sailors in a ship that must attack and pillage other ships and towns in an environment composed of three islands and the sea areas between them (Schell & Shochet, 2001). Figure 12 shows a schematic diagram of the attraction. This attraction has been conceived for up to four users acting as the captain and three gunners (Figure 12a).

The attraction is based on a motion platform that defines part of the deck of the pirate ship (Figure 12b). The user that acts as the captain leads the ship with a physical helm (Figure 12c). The other three users, the gunners, may shoot at other ships or towns with six physical cannons that are placed three on each side of the ship (Figure 12d). The users see the environment through LCD shutter glasses on the stereo screens around them (Figure 12e). They see the images of the environment in a first person view scheme from the ship (Figure 12h) such that part of the ship is virtual within the images (Figure 12f).

The VS has a complex physical interface, the physical ship, which is composed of several elements (both input

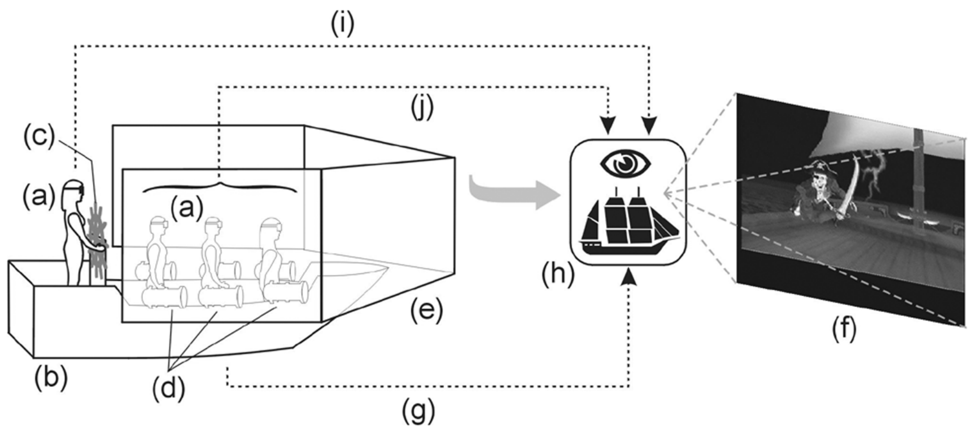

Figure 12. *Schematic diagram of the "Pirates of the Caribbean: Battle for Buccaneer Gold" attraction (see text for details).*

and output), namely: the motion platform, the helm, the cannons, and the screens. It also has several elements that form the logical interface, the virtual ship (Figure 12h). These logical interface elements are the hull of the ship, the cannonballs, and the point of view centered on the ship. There are a set of mappings defined between the elements of the physical interface and those of the logical interface. For example, there is a mapping between the helm of the physical ship and the rudder of the virtual ship (Figure 12i) that relates equivalent rotations. There is also another mapping between the physical cannons and the invisible virtual cannons (Figure 12j), also related by rotations to orient the shots properly. Finally, there is a mapping between the hull of the virtual ship and the motion platform (Figure 12g) such that any movement of the virtual ship is translated to the physical ship. Apart from this, we must also consider the behaviors that are defined for the logical interface. Some of these behaviors are the collisions of the hull of the virtual ship against the waves of the virtual sea; the shooting of the virtual cannonballs from the virtual cannons and how these cannonballs affect other ships; and the action of the rudder against the virtual sea to change direction of the ship.

The success of the design comes from the fact that, although the VS is a sophisticated structure, it is very clearly defined and identified: "the ship." This is a crucial point because the users immediately understand that they all constitute a single unity within the game. Although they are four distinct users they are working together to be a single unity within the experience. The captain does not lead the ship independently from the gunners. Moreover, the change in direction of the ship affects the viewpoint of all the users. Any treasures won are counted on a common score, and so forth. Therefore, the success of this application is that the mental model that the users have is correctly matched to the VS of the experience.

We can clearly see how the "Hercules" application has not followed a consistent design of the VS for the users and therefore the experience is incoherent for them. On the other hand, the "Pirates" application is extremely solid in its conception and therefore mediates the experience very well to the users.

Someone may find these examples to be too trivial; that they may be analyzed without the need of the VS. This is in fact true; however, it is also a fact that these errors keep appearing too often, even in important design and development teams. We believe that the sole fact of understanding and bearing in mind the idea and structure of the VS and how it mediates the experience of the user should already help to minimize these errors. Of course there is now a need to establish all the possible interrelations between the components of the VS to fully understand how they mediate the experience and, from there, to define a detailed framework that can allow us to analyze applications and give us the elementary steps to design good experiences.

9 Conclusions

In this paper, we have presented a preliminary model to describe a VR experience with respect to how the user is confronted to it, how she perceives it and understands it, and how this can help, not only in formalizing the properties of such experiences, but also in designing new experiences. This has been done through explicit differentiation of the terms virtual environment (VE), which is the static definition of structures, and virtual reality (VR), which is the actual experience of the VE when this latter is put in action and related to the user. The model then defines the key element of the user-experience relationship: the VS, a high level element that fully links the user and the VE in a VR experience. This VS is composed of:

- the *logical interface,*
- the *physical interface,*
- the *mappings* between them, and
- the *behaviors* of the whole set (especially of the logical interface) with respect to the VE,

thus generating the experience.

We have briefly given a possible initial framework although it must still be elaborated and detailed in order to become a useful design and analysis tool. This ongoing research should also lead us to finally understand the underlying processes that control the mediation of the experience to the user. However, it already gives a clearer description of user experience, unlinking it from any specific VR technology and not restricting these experiences to the area of simulation applications, therefore leaving open the full range of possible experiences. It must also be analyzed theoretically to see how it may help in clarifying the specific properties of VR as a communication medium. This should lead us further in the process towards obtaining a solid model for a virtual reality experience.

Acknowledgments

We would like to thank the Audiovisual Institute of the Universitat Pompeu Fabra for providing the fruitful and rich environment that has made possible the research that has led to the present paper.

References

Biocca, F., & Levy, M. R. (1995). Virtual reality as a communication system. In F. Biocca & M. R. Levy (Eds.), *Communication in the age of virtual reality* (pp. 15–31). Hillsdale, NJ: Erlbaum.

Biocca, F., Kim, T., & Levy, M. R. (1995). The vision of virtual reality. In F. Biocca & M. R. Levy (Eds.), *Communication in the age of virtual reality* (pp. 3–14). Hillsdale, NJ: Erlbaum.

Bowman, D., Kruijff, E., LaViola, J., & Poupyrev, I. (2001). An introduction to 3-D user interface design. *Presence: Teleoperators and Virtual Environments, 10*(1), 75–95.

Brown, P. (1994). The ethics and aesthetics of the image interface. *ACM SIGGRAPH Computer Graphics, 28*(1), 28–30.

Carlsson, C., & Hagsand, O. (1993). DIVE—A platform for multi-user virtual environments. *Computers and Graphics, 17*(6). URL: http://www.sics.se/dive/. Accessed April 2006.

DisneyQuest. (1998). Walt Disney World, Orlando, FL. http://www.disneyquest.com/. Accessed April 2006.

Ellis, S. R. (1991). Nature and origins of virtual environments: A bibliographical essay. *Computing Systems in Engineering, 2*(4), 321–347.

Ellis, S. R. (1993). Prologue. In *Pictorial Communication in Virtual and Real Environments,* 2nd Edition (pp. 3–11). London: Taylor & Francis Ltd.

Ellis, S. R. (1996). Presence of mind. A reaction to Thomas Sheridan's "Further musings on the psychophysics of presence." *Presence: Teleoperators and Virtual Environments, 5*(2), 247–259.

Clements, R. (Director). (1997). *Hercules* [Motion Picture]. United States: Walt Disney Pictures.

Krueger, M. (1985). Videoplace—An artificial reality. *ACM Conference on Human Factors in Computing Systems,* 35–40.

Latta, J. N., & Oberg, D. J. (1994). A conceptual virtual reality model. In *IEEE Computer Graphics & Applications, 14,* January, pp. 23–29.

Mine, M. (2002). Towards virtual reality for the masses: Ten years of research at Disney's VR studio. In *Proceedings of the Workshop on Virtual Environments 2002, EGVR02.*

Minsky, M. (1980). Telepresence. *OMNI 2,* 45–51.

Norman, D. A. (1988). *The design of everyday things.* New York: Basic Books.

Parés, N., & Parés, R. (1993). Galeria virtual. In *Proceedings of the First Eurographics Workshop on Virtual Environments.*

Parés, N., & Parés, R. (1995). Galeria virtual: A platform for non-limitation in contemporary art. In R. A. Earnshaw, J. A. Vince, & H. Jones. (Eds.), *Virtual reality applications.* London: Academic Press.

Pirates of the Caribbean. (1967). Disneyland, Anaheim, CA. URL: http://www.disneyland.com. Accessed April 2006.

Rodrigo, M. (1995). Los modelos de la comunicación (Communication Models), 2nd edition, Madrid: Editorial Tecnos. (in Spanish).

Schell, J., & Shochet, J. (2001). Designing interactive theme park rides. *IEEE Computer Graphics and Applications, 21*(4), 11–13.

Sheridan, T. B. (1996). Further musings on the psychophysics of presence. *Presence: Teleoperators and Virtual Environments, 1*(1), 120–126.

Slater, M. (2003). A note on presence terminology. *Presence-Connect, 3.* http://presence.cs.ucl.ac.uk/presenceconnect/.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 6*(6), 603–616.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication, 42,* 73–93.

Stone, V. E. (1993). Social interaction and social development in virtual environments. *Presence: Teleoperators and Virtual Environments, 2*(2), 153–161.

Vincent, V. J. (1993). The Mandala virtual reality system: The vivid group. *Proceedings of the Third Annual Virtual Reality Conference and Exhibition on VR Becomes a Business,* 167–170.

Waterworth, J., & Waterworth, E. (2003). The core of presence: Presence as perceptual illusion. *Presence-Connect, 3.* http://presence.cs.ucl.ac.uk/presence connect/.

Whicker, M. L., & Sigelman, L. (1991). *Computer simulation applications; An introduction.* Newbury Park, CA: Sage.

WorldToolKit. (1999). Sense8 Co., WorldToolKit Release 9.