

Chapter 1

Social Interaction in Virtual Environments: Key Issues, Common Themes, and a Framework for Research

Ralph Schroeder

In this chapter, I will give an overview of some central issues in research on shared virtual environments (VEs) – including “presence”, “copresence”, communication, and small and large group dynamics - across a range of virtual reality (VR) technologies and different conditions under which they are used. I will discuss different studies of the interplay between technological systems and their social implications, and how sociological insights about interaction in the real world can be brought to bear on interaction in VEs. Finally, I will argue that making links between different areas of research can lead to a better understanding of social interaction in VEs.

1.1 Background

In the early 1990s, the dominant image of VR, and what most laboratories and developers focused on, was of single-user head-mounted display (HMD) systems [1, 2]. Nowadays, there is a range of technologies, from expensive and immersive projection technology (IPT) or CAVE-type [3] room-size VR systems in which the environment is projected onto several walls, via HMDs, to free VR software that runs on desktop personal computers (PCs).

Only since the mid-1990s, with the popularity of the Internet, has it become feasible to link many users simultaneously in shared or multi-user VEs. Today there are dozens of internet-based VEs that can be run on PCs, and in which hundred of thousands of participants have created virtual social institutions such as shopping malls, churches, museums, classrooms, and the like. There are also dozens of trial systems being developed in computer science laboratories around the world that aim to develop shared VEs for a variety of purposes; among others,

virtual business meetings, scientific co-visualization, virtual therapy, and entertainment. These experimental systems often make use of more complex technological systems such as high-end computer graphics workstations, HMDs, and a host of other display systems, input/output devices, and computer graphics software.

Despite this proliferation of technologies, and although the word “virtual” has come to be used in lots of different ways, there is nevertheless a core area of research on multi-user VR/VEs that most researchers would recognize. I have previously defined virtual reality technology as “a computer-generated display that allows or compels the user (or users) to have a feeling of being present in an environment other than the one they are actually in and to interact with that environment” [1]. This definition is close to that of most researchers in the field, and it is also grounded in a particular understanding of the social implications of new technologies. *Shared* VR technology, or *shared* VEs, can therefore be defined as VR systems in which users can also experience other participants as being present in the environment and interacting with them.

1.2 Methods and Approaches

There are two main methods for studying social interaction in VEs. Experimental methods typically make use of “purpose-built” environments to study a controlled set of conditions, whereas qualitative methods, such as participant observation, are often used to study “naturalistic” settings. This need not be so: Chapter 12 in this volume examines a “naturalistic” VE setting in a quantitative way. Nevertheless, most studies are either based on short and controlled trials or on longer-term observations of what people do in ongoing VE settings.

Some research areas lend themselves more to one or other method. It is difficult to envisage how it would be feasible to study how people build complex virtual settlements by means of experimental methods, or under controlled conditions – though chapter 5, *30 Days in Active Worlds*, comes close to being an experiment of how people build in VEs. Or again, experimental results about how people interact in short collaborative tasks (see, for example, Chapters 9 and 10) – relating to how people work together, for example – might not apply if the setting was a more “naturalistic” one, or one where the subjects were not under the experimenter’s gaze or influenced by their instructions. It is difficult, however, to make trials natural, or to carry them out over longer periods.

In our studies, and in the chapters in this volume, a variety of methods have been employed, including various kinds of experimental studies and forms of participant observation. The latter has involved spending a long time, especially in *Active Worlds* (AW) – one of the most interesting online VEs which is discussed in several chapters – taking detailed notes on particular phenomena that are of interest, or conducting semi-structured online interviews with users. Experimental studies often vary the conditions – say, with different VR systems – and compare the results. These studies also often make use of questionnaires to get the responses of the “subjects” and sometimes use audio or video recordings.

These different methods point to the variety of approaches in the study of shared VEs: on the one side, there is the more technically oriented literature which often comes from researchers in computer science departments and covers VR technology, collaborative virtual environments, human factors, computer-supported cooperative work, and the like. On the other side, there is the social science literature around MUDs (Multi-user Dungeons or Dimensions), identity on the Internet, or new media and society [4, 5, 6]. In this case, researchers often come from sociology or media/communications studies departments.

Clearly, it will be useful to continue with a variety of methods, but it is worth making some brief comments. One is that the questions in studying social life in virtual environments are still emerging. As we shall see, the findings of studies are often from initial trials or from early uses of systems. Another is that a number of questions about research ethics and methods remain, as we found, for example, in our study of a church service in an online shared VE: should the settings and informants be treated in the same way as in the real world? Are informants who are only encountered virtually reliable? [7]. Finally, it can be hoped that findings using different methods will have the positive result that they complement each other, though it remains to be seen how well-integrated research on social interaction in VEs will become. With this, we can turn to the key substantive issues in shared VEs.

1.3 Presence and Copresence

“Presence” is a term that will be familiar to VR/VE researchers, but it will not be familiar to those outside of this research community [8]. VR technology, as indicated by the definition given earlier, is about “being there”: presence is therefore partly to do with the technology, and partly to do with the participants’ state of mind. A recent overview of research [9] discusses several concepts and ways of measuring “presence”. This overview also covers some commonly used indicators of presence, such as “immersion” and “involvement” in the environment. A further debate that they review is between “subjective” measures of presence, which are often obtained by means of questionnaires, as against “objective” measures, which entail, for example, the timing of task performance or heart-rate measurements.

Much of the experimental research on presence to date has been on immersive systems (IPT systems, HMDs), or comparisons with desktop systems. This research includes several studies where participants carry out a task first in one system and then in the other, or where participants using one type of system collaborate in the VE with participants using a different type of system. These studies often show that participants experience a greater sense of presence in more immersive systems than in less immersive ones.

However, it is important to broaden this discussion. Research with immersive systems typically involves short and controlled trials and a particular task. Yet users of AW, who typically spend a long time in this online shared VE, often

without a particular task except socializing, surely also have a sense of “presence”. They may be less focused on their activity, and have a system with poorer 3D graphics – and since they may also be using the AW VE in a fairly routine, on again/off again manner – they may not think of themselves as being “present” in the same way as users of immersive systems. Nevertheless, they clearly have a sense of “being there”, even if “presence” is not an issue for them. Still, “presence” is typically discussed in the context of immersive rather than desktop systems [9].

The same point can be made differently: in presence research that has involved short experimental trials, users will typically answer questions about presence on Likert-type scales. However, if, for example, users regularly spent time in a highly immersive VE such as an IPT system, would they respond to such questions about presence differently? We can see then that the immersiveness of the technology per se may only be one dimension of presence, and the “mundaneity” of use may be another.

Presence thus depends on a variety of factors, including the task, the VE, and in shared VEs, on copresence (as we shall see in a moment), and these factors will often outweigh the technology in affecting presence. Ultimately, if we want to measure different degrees of presence in VEs objectively, we may only be able to do so, as Ellis has argued [10], by measuring different variables and *against each other* (Ellis uses the notion of “equivalence classes”). In other words, it may be necessary to create different VE scenarios that are comparable, and systematically measure presence in one against the other, not only in terms of performance as Ellis suggests, but also in the light of other variables. Another way to study presence will be to compare presence and interaction in different types of VEs with equivalent *real* scenarios [11, 12], as well as with other mediated environments. A combination of these methods will ultimately lead to a more comprehensive and thorough understanding of presence.

Note that presence does not depend on the fidelity or “realism” of the VE: a “fantastical” or “abstract” VE can also provide a sense of “being there” for the user. Moreover, a number of studies have shown that presence does not necessarily increase task performance [9]. The reason for this may be that users need to divide their attention between the environment and the task in a situation where both are highly engaging.

This brings us to “copresence”: presence, or “being there”, and copresence, a sense of “being there together”, are bound to be closely related. Again, we can initially widen the discussion instead of focusing exclusively on the results of experimental studies. For other media, issues similar to “copresence” are often discussed under the rubrics of “social presence” or “media richness” (see Chapter 10 in this volume and [6] for reviews). Shared VEs are a *rich* medium in the sense that they allow people to interact via several senses. In the case of most of the VEs treated in this volume (with the exception of haptic VEs, see Chapter 10), they allow people to interact via audio/text and via a 3D visual environment. This sets shared VEs apart from telephony, video conferencing, and other media of communication – though whether they are similarly useful or enjoyable media remains to be seen. Nevertheless, it is a popular reaction to shared VEs, especially among novice users, to comment on how lifelike they are.

The richness of this medium has been demonstrated in various ways. Slater and Steed (Chapter 9), for example, have shown that people have a very strong reaction to others, even if they are merely computer-generated agents. And as their acting trials show, forms of interaction which require sophisticated social cues – such as acting together – are feasible in shared VEs. Another example is our trial, carried out jointly with Steed and Slater, with networked IPT systems, in which we demonstrated that certain tasks can be carried out remotely as if “being there together” [13]. In this trial, we linked two IPT systems to allow two people – one in London, one in Gothenburg – to collaborate on solving a Rubik’s cube-type puzzle (see Figure 1 in Chapter 11). We demonstrated, by comparing this with the equivalent task carried out face-to-face with cardboard boxes, that such a highly spatial and collaborative task can be done just as effectively in networked VEs as in a real face-to-face setting.

Some elements of shared VEs, on the other hand, detract from the richness of the medium: one is that many social cues are missing. For example, communication in shared VEs is often via voice, but many bodily cues are missing. And although non-verbal communication is sometimes used, some studies suggest that it is not used as much as in face-to-face interaction – even if many of the rules of non-verbal communication, such as turning your gaze in the direction of your intended audience, *are* adhered to (see Chapters 2 and 12).

Many rules that govern copresence will be affected by technology, and here systematic comparison between different technologies will be useful. This is an obvious point, but it deserves restating here in order to stress that this is not just a question of “high-tech” versus “low-tech”, or highly immersive systems versus desktop systems. Rather, as Chapters 5 and 11 show, there are many features of “low-tech” desktop systems, such as access privileges or technology for the appearance of the environment, that have consequences for “copresence”.

Nonetheless, shared VEs often combine a high degree of presence with a high degree of copresence because the sense of being in another place and of being there with another person reinforce each other. It may seem self-evident that presence and copresence *should* go together, but although we have some studies that point in this direction (Chapter 9, [13, 14]), we lack research, again, of the type suggested by Ellis, whereby a number of comparable settings are studied against each other.

It also seems possible that the effect of copresence may “wear off” as the novelty of the medium wears off. At the same time, it seems likely that copresence will increase with the degree to which copresent users establish strong relationships in VEs. The first is something that shared VE researchers will know from their own experience, the second is clear from interviews with long-term users [15].

As in the case of presence then, it appears that users are able to cope well with the absence of certain features of the real world or of face-to-face copresence – while they also need other features. As Buescher et al. [16] have argued, participants in shared VEs need at the very least a reciprocity of perspectives to make sense of each other’s actions. Reciprocity is thus one of the most elementary building blocks of social interaction, and from an analysis of this reciprocity, and how it influences how copresent participants focus on or turn their attention away

from others, it will be possible to build up a picture of more complex forms of interaction.

1.4 Communication

To study communication in shared VEs will require a combination of perspectives, including social psychology, sociological analysis of interaction, and communications studies approaches to different media. Shared VEs typically make use of either audio or text-communication plus non-verbal communication.

First, I should explain briefly why text-based communication is included in a discussion of shared VEs. In a sense, text-based communication (as, for example, in AW) is not VR because it does not enhance – but rather detracts from – the sense of presence and copresence. Nevertheless, the reasons for including text-based communication here are: first, that it is widely used in large-scale internet-based VEs; second, there is an extensive research literature on text-based computer-mediated-communication (CMC) which, as several chapters demonstrate (especially Chapter 2 and 11), can be usefully brought to bear on social interaction in VEs; and, finally, the study of VEs not only benefits from comparisons with other forms of CMC, but also from comparing different modalities *inside* VEs (see, for example, Chapter 10). Arguably, VEs will never provide completely “realistic” ways of interacting or communicating with others because a number of features of face-to-face interaction will always be lacking. It is therefore instructive to compare different modes of communication in VEs, for example text with voice (Chapters 2 and 10) or with face-to-face communication [17].

In relation to shared VEs that support audio communication, one finding that has emerged again and again (see the results of the COVEN trial [18], and Chapter 7 in this volume) is that the quality of the audio communication can be a major obstacle to collaboration and fluid interaction. It can be anticipated that this technical problem will be overcome, but there is also an implication for the design of shared VEs here: there is little point in developing a technologically sophisticated or visually complex shared VE unless the audio communication works well, since this is critical for effective or enjoyable interaction.

Some evidence for low media richness or low social presence in VEs is that people do not use non-verbal communication as much as in face-to-face interaction – as mentioned earlier. This does not mean that they do not use their bodies to communicate (see Chapter 3 for examples). Nevertheless, the dearth of non-verbal communication needs to be set against the observation that users seem to be able to adjust easily to communication in shared VEs.

Analyzing communication in VEs can also take place on different levels: for large groups like a population in AW, we can examine the use of language by looking at the encounters between different national languages, for example, or at greetings, or at the number of words per contribution, and how this might differ from “real world” conversation structure (Chapter 2, [17, 19]). In small groups, we can analyze whether the VR system used makes a difference to who dominates

verbally in a task scenario – for example, in carrying out a spatial task which requires a lot of communication and where participants are using different systems [13, 14], one of which is better suited to the spatial task. Again, both types of study will be useful since effects such as increased or diminished equality due to technology will operate across both large and small groups (see Chapter 11).

1.5 Avatar Appearance and the Appearance of the Environment

One interesting “lesson learned” by the developers of Microsoft’s V-Chat system (see Chapter 6) is that people would like to be able to have greater control over their own avatar representation or have input into its design. This demand from the users is also supported in Chapter 3, and – as I found out in an interview with the developers of the AW system (see [20]) – it is also the most common request that AW users have.

It is interesting to note that meeting this demand currently presents several technological challenges: one is to provide the user with the tools to create their own custom avatar. Another is to do with network capabilities: in shared VEs, should each avatar representation permanently reside on each of the other users’ computers? This would create memory problems if there were hundreds or thousands of unique users. Or, should each new avatar only be downloaded when it is used? This would avoid the storage problem on each computer, but it would require lots of bandwidth. A related issue here is the complexity of, or the amount of data required by, each avatar. Finally, a few customized avatars may not be a problem, as in the case of small groups, but in larger populations there will continue to be a trade-off between unique and complex avatars and technological capabilities. As it stands, therefore, it is only possible to have a small number of custom avatars.

No doubt technical solutions will make progress here. It is also possible to anticipate that avatars will feature a mixture of computer-generated representations and real-time video images of users [21] – so that avatars will range from cartoon-like, as they are now, to very realistic. Chapter 9 provides some images of avatars that are quite realistic. But as Cheng, Farnham and Stone have found (Chapter 6), users may want avatars that are neither too abstract nor too realistic. It is therefore too early to say how much avatar customization will in fact be demanded by users in systems where they are given a choice.

Perhaps a mixture of “off-the-shelf” or ready-made avatars and customized solutions will emerge in VEs. In the meantime, several chapters in this volume (Chapters 3, 8, 9) provide some indications of the effects of avatar appearance on social interaction. My point, again, is that what needs to be considered here are not just the effects of avatar interaction on individual encounters, but also issues such as the influence of the persistence of avatar appearance in different conditions: for example, what kind of persistence do users need in order to recognize each other over repeated encounters? And what kind of diversity of avatar appearances is

needed within both small and large groups for participants to be able to distinguish one another – and what diversity can they cope with?

For the features of the landscape and the built environment in shared VEs too, there are points of overlap between online shared VEs with many users and immersive VEs for small groups. Chapter 5, for example, points out that there is a build-and-abandon attitude in AW. This applies to immersive VEs too, since many VEs are developed and built, often at great cost in terms of labor, but are only used for a few demonstrations – after which they are abandoned and never used again. Again, this observation has implications for the design of VEs.

Environments with large populations often have extensive and varied landscapes, such as the hundreds of worlds in AW. Perhaps the easiest way to make this point for those who are not familiar with AW is to say that it would take many days to see the various sights and to become familiar with the social milieus that can be found in the many worlds that have been built in AW. And again, AW is only one – though perhaps the most interesting because it has largely been created by users – among several internet-based social VEs.

Another feature that should be mentioned is the mixture in the environments of elements that imitate the real world as against those that depart from it; or, real versus imaginary VEs. Examples of “realism” include the way that the layout of densely populated areas imitates real world cities, the resemblance of many buildings in AW to real world buildings, and the furnishing of many houses with chairs and tables (which serve no function apart from decoration or orientation); examples of the “imaginariness” in AW are the frequent use of all-glass transparent buildings, buildings which imitate science fiction or which are built in the sky, and objects like waterfalls or flames in unlikely places.

In experimental studies, the appearance of the VE is typically related to the topic under investigation: visualization, collaboration, acting rehearsals, etc. There are also some highly realistic environments for training and games; military simulations and internet-based games like Quake and Doom are good examples of the latter (though they fall outside the definition of VR given earlier). These may have a higher degree of “realism”, but they are often restricted to a particular functionality: the user must follow certain rules (in a game) or manipulate the environment by means of certain tools or weapons.

Shared VEs that are used in experimental trials will often be more restricted in scope and more abstract. Online VEs, as mentioned, are very extensive and mix fantasy and realism. And although there are studies [22, 23] of the geography of (mainly online) shared VEs, we need a more comprehensive classification of the appearance of VEs. Such a classification would be useful because it would allow us to relate their appearance to how users interact with them: what features must the environment have in order to enable particular types of social interaction? This is an issue which goes beyond joint navigation or wayfinding [24], and there is often a mismatch here, especially in that environments are often too complex for users’ needs.

It may seem obvious, again, in taking a broader view, that the appearance of the environment will affect not only navigability (the issue that has been most studied so far), but also how users interact with each other. As we shall see below, this is an issue that can be framed in terms of how much the appearance of the

environment detracts from – or allows the user to focus on – interaction with other users.

1.6 The Dynamic of Small and Large Groups

Apart from copresence and communication, we also need to analyze social interaction *per se*. Several studies in this volume show that – in small groups and in large ones – users, to a large extent, follow the conventions of the real world. These conventions include keeping their distance from each other, turning to face their conversation partners, and so on (Chapters 2, 8, 9, 12). It is equally clear from a number of studies, however, that users do not follow these conventions in other respects: for example, they do not use gestures very much, they more readily abandon and destroy buildings (Chapters 4, 5), or they treat a person with more powerful VR technology as the leader, even though they don't do this in the equivalent real situation [11].

This is an area where so far few links have been made between the rules governing small and large groups (but see [25]). Studies of shared VEs, and of CMC generally, tend to focus either on small groups of up to three or four, or they study large groups or populations in shared VEs or CMC. Some of the intermediate levels have been analyzed for other forms of CMC – such as the use of email tools in organizations – but this has not been done for VEs. In shared VEs, perhaps the closest we can come to this intermediate level are the inhabited TV trials [26]. It is interesting to note that in these settings, where there were dozens of participants, one limitation that became apparent was that not many could actively participate, and a divide emerged between active participants and onlookers. In other words, in shared VEs, as in the real world, the focus of attention needs to be concentrated on a few members, and it is difficult for people to participate actively in a large group event.

It can be noted in passing that this also points to a limitation on the notion of “interactivity”, which is often used in discussions of VEs and other electronic media. As we can see, in shared VEs, even if the user possibly experiences more interactivity than in most other new media technologies, nevertheless, the interactivity *with others* inside the VEs is subject to similar constraints as in real life – perhaps even greater ones.

The gap between the study of small groups and larger populations is also characteristic of social science as a whole, where micro- and macro- are not well integrated (though see [27] and [28] for attempts in this direction). Nevertheless, there are crucial links here: for example, whether one's role in a small group is recognized as much in a VE as in real settings (as some studies have analyzed) will also be significant for larger groups or populations. And vice versa – whether leadership roles or hierarchical roles are generally acknowledged in large groups in shared VEs as in the real world, or if there are status equalization effects (see Chapter 11) will also translate into small group behavior. These links between small and large groups apply to a range of issues, and eventually the study of shared VEs will have to bring them into a comprehensive framework of analysis.

1.7 Relation to Offline Behavior and to Other Media

One current limitation of the study of shared VEs is that we know little about how online behaviors affect users' behavior offline. This is partly because shared VEs are often studied in the context of short task-related trials. Even where there is research on longer-term shared VE settings, such as online ones, there are no studies yet which relate this to the real-life context of users: in the case of "social" VEs, which are the most widely used, there are no studies which have related "online" and "offline" life, though there are some studies which have related the two for *text-based* social MUDs [29, 30]. And although there has been an overview of after-effects research, mainly for immersive systems [31], this has mostly been concerned with problems such as short-term disorientation or nausea, rather than with the effects of interacting with others in the VE or with the real-life contexts of users.

I have argued elsewhere [25] that, in the end, the study of social interaction in VEs needs to be integrated with the study of the uses of other communications media and how these media, including VEs, fit into our everyday lives. By comparing shared VEs with other forms of CMC and other media, and combining this with studies which compare virtual versus real interaction during short trials and how this affects social interaction (such as leadership and embarrassment, see Chapter 9), we may eventually be able to relate virtual and real interaction more systematically.

1.8 A Framework for Research: Frames, Focus, Roles, and Networks in VEs

Identifying key issues and common themes provides the backdrop for sketching a framework that brings together the various facets of interaction in shared VEs. One way to bring insights from the social sciences to bear on shared VEs is to start with Goffman's ideas about the "frames" of social interaction [32, 33]. For Goffman, frames are the stages on which we play out our social roles. However, in shared VEs, the way we act and interact with others is technologically mediated. Thus VEs have a different kind of "bandwidth" from real world frames for presenting the self to the other. When we enter a VE, a shift in the "frame" takes place, and the bandwidths (in the non-technical sense) of different types of VE vary a great deal between, say, highly immersive and non-immersive VEs (see Figure 1.1).

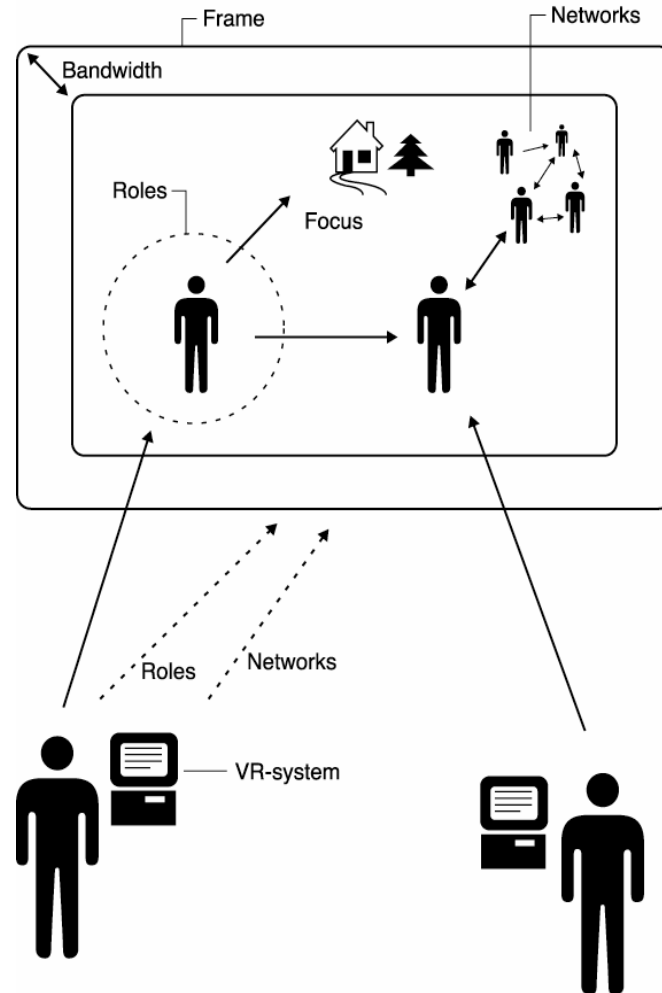


Figure 1.1 Frames, focus, roles and networks.

“Bandwidth” in this case is not the same as “media richness”, “communication modality”, or “social presence”. Instead, using the notion of frames with different bandwidths in a VE allows us to apply the rules of face-to-face interaction to interaction in different shared VE settings: that is, different VR systems or types of VEs provide different frames for our encounters. And, to anticipate, if this applies to individual encounters and how we present ourselves to each other, it will also apply to larger groups.

Before we proceed further with this framework, two asides are necessary: the first is that although Goffman is often interpreted as arguing that “it all depends on context”, or that social interaction is “relative” to the particular “frame” in which it takes place, he can also more accurately be regarded as advocating the systematic study of frames whereby generalizations can be made *across* different frames or contexts [28]. In other words, Goffman’s ideas can be seen as part of an objective social science which applies not just to individual frames of interaction, and which can also be incorporated into a more macro-analysis of social structures. Second, while Goffman did not apply his ideas to communications media, they can nevertheless be extended to provide powerful insights into the social implications of electronic media, although this is beyond the scope of this chapter (but see [34]).

For Goffman, the next step in the analysis is to look for the focus of attention in the social interaction or in the encounter between people. In shared VEs, as we saw earlier, the degree of focus (and distraction) that is possible relates to presence and copresence – or here, to the bandwidth of the frame. So, for example, some VEs are (initially) so visually rich as to overwhelm the user. In other VEs, the environment may be abstract or information “poor”, leaving the user to focus on the task or on the interaction with the other participant(s). In other words, here the focus is on what people *do* together. In this way, the question becomes not so much “how present do users feel”, but where is their attention focused in the VE? (Benford and colleagues have also used the notion of “focus” in VEs [35], but they use it to analyze spatial orientation, not social interaction).

In certain small group studies, such as our Rubik’s cube trial (see Figure 1 in Chapter 11 and [12]), there is a high task focus and the focus on interpersonal relations is secondary. In other trials, such as the acting trial reported by Slater and Steed in Chapter 9 or in Blascovich’s chapter (Chapter 8), there is a high interpersonal focus because the interaction revolves around interaction with others rather than a shared practical task. We can also see a high focus on interpersonal relations in the close-knit groups that are described in the chapters about online VEs (see Chapters 4 and 5, [15]) where a lot of attention is paid to the way participants present themselves to others and to how the rules of relationships are followed.

“Focus of attention” applies not only to how the user perceives or engages with the environment and other users, but also outside it – how much distraction there is from the frame of the VE. An HMD system, for example, will almost completely “shut out” the world around the user, though it has often been noted that wires and other obstacles may distract the user. Or again, an IPT-type system, which may provide a greater sense of presence in terms of “place” than an HMD system, may leave the user with a peripheral sense of others being there in the real world (outside the walls of the IPT-type system), and therefore diminish the copresence with those inside the VE. Similarly with desktop systems: users may have a high degree of copresence if, for example, they are engaged in a highly engaging spatial task, or participating in an online religious service – both of which involve a common focus of attention. However, this sense of co-presence can be weakened if they split their attention between others copresent in the VE and another person sitting beside them in the real world.

The difference between this and other frames, which makes the very definition of VR/VE in terms of “being there” so important, is that VE frames are entirely technologically mediated. However, while frames should be analyzed in terms of technological mediation, the analysis of interaction *within* frames proposed here is much like the analysis of real world interaction, unlike other theories of new media which focus on “interactivity” or “media effects” and the like. In other words, this framework does not treat shared VEs differently from real world interaction, except in aiming to compare and contrast CMC with face-to-face interaction, and putting the use of shared VEs into the larger context of our uses of CMC and other new media in society.

Another key aspect of applying frames to shared VEs is how much people have become used to VEs. As Chapter 6 shows, regular users navigate less. Perhaps this indicates that they have become more focused on interpersonal relations rather than on moving around in the environment. Axelsson and Schroeder similarly found this when interviewing regular users of AW [15]; their involvement in the environment depends on how much they have built and routinely interacted with others. This also became clear over the course of hourly sessions in AW (see Chapter 7); during these sessions with their variety of activities – building together, exploring, making presentations, planning, and the like – attention was unevenly divided between the environment, the others, the task, and, peripherally, the real world.

The key variable within the frame is therefore the focus of attention – on the co-present others, on the task or interaction, and on the environment. The frame, its bandwidth, and our focus in it – what I discussed earlier under the headings of presence, copresence and modality of communication – thus shape how we experience the VE as a place and how we engage with others. There will continue to be various approaches to studying presence and its facets – immersion, “being there”, etc. What I am suggesting is that a person’s presence in *shared* VEs can be seen as part of their *interaction* with others, which includes how we present ourselves to others and encounter them in small groups.

Communication, and especially the modality of communication, can thus also be incorporated within the framework of frames and bandwidth: different types of shared VEs will provide different opportunities and constraints for presenting ourselves to, or communicating with, others. Some examples are described in Chapter 2. However, we can also examine language or communication in relation to the focus of attention in small groups: who takes a dominant role in communication in relation to a certain task [11, 13], or how task related or non-task related (socializing) the conversation is, and so on. And we can compare different modalities of communication, as Sallnäs does (Chapter 10). In relation to text-based communication in shared VEs, it is clear that there are major differences in form and content from VEs with audio or from face-to-face communication: in text-based VEs, there is much more focus on addressing each other and gathering contextual information, shorter exchanges, etc. (Chapter 12, [17]).

From encounters between individuals, where the self is presented to the other, we develop different roles vis-à-vis others in different circumstances, and thus also take on different roles in different networks of relationships. Before moving on to the discussion of roles, however, it is important to note that our roles in shared VEs relate to focus: focusing our attention on several people in a VE can be

burdensome, and it seems that this is often an obstacle in small groups in shared VEs, especially in comparison with real settings. This is partly to do with the restricted field of view, and partly with the absence of social cues in VEs. Put differently, it is difficult to experience the “cocktail party” effect in shared VEs whereby, in the real world, we can follow a conversation across the room. Yet not being able to cope with many simultaneous complex impressions will also leave room to focus on other aspects of our interaction with others, as we have already seen in a number of examples.

Roles also depend on the setting. In the “social” VEs that are described in several chapters, there is a certain “frisson” for novices who encounter other people in the form of avatars for the first time, and they may experiment with how they present themselves and which rules they break or follow [30]. Nevertheless, participants will also compensate for the absence of social cues and for well-defined roles by presenting lots of information about themselves (name, age, sex, location) and gathering as much information as they can about others. This is a form of interaction to which participants quickly adapt. And over longer periods, we have found [15], as has Schiano [36], that participants generally maintain stable roles (or “identities”) and increasingly adhere to the norms that they have come to share with others. (It can be mentioned in passing here that the novelty or familiarity of VEs also ties the frame of the VE to its offline context: online VEs, where participants invest more or less in their role or their “online persona”, are an obvious example.)

Interestingly, a number of studies [13, 14] show that role differentiation or a “division of labor” can take place “automatically” in shared VEs because of the different technological capabilities of the participants – *even* when they are not aware of the difference between the systems they are using. So, for example, in our Rubik’s cube trial (see Figure 11.1 in Chapter 11), the person in the immersive VE concentrated on the spatial task while the person on the desktop system stood back and verbally supervised.

However, the strength or weakness of the role – for example, leadership in a group [11] – is not just a product of the particular encounter or situation, but also depends on how strongly roles are shaped or defined in the shared VE as a whole; in other words, how pronounced the system of roles or of stratification is (Chapter 11, [20]), and this will carry over from small groups into larger ones, and vice versa.

Frames apply to individual and small-scale encounters, but they also apply to groups, with virtual meeting places as the stages for larger gatherings. From encounters, where participants develop roles *vis-à-vis* each other, share perspectives, and engage in relationships of reciprocity, we can thus move to analyzing social networks. Yet even here, for shared VEs, it is clear that the frame and bandwidth will dictate the density (or lack of density) in larger online networks.

The notion of social networks is particularly useful in this case since all shared VEs involve technological networks that create new relationships between people. From roles, which are suitable for the study of encounters in small groups, we can thus move to larger groups and begin to examine phenomena such as the differentiation of roles within groups, the division of tasks (or of labor) (see above)

and the exchange of resources (see Chapters 3 and 4). A number of larger social phenomena described in these chapters and elsewhere – building, social conventions, etc. (Chapter 2, [7, 15, 21]) – apply both to small and large groups or networks.

Much here depends, as these studies also show, on how much users have come to know each other and how much they have become involved in the VE, including shaping and adhering to its norms and helping to build the VE [22]. And again, networks tie online to offline behavior, as when we can map the online relations onto real world relationships, say, in collaborative groups which meet both on- and offline, or in the offline conventions that AW users have in addition to their online meetings.

Apart from the links in larger groups, the study of social interaction will want to examine online social structures in various ways – as in the social world at large. However, for shared VEs, it is likely that networks will always play a key part in the analysis since shared VEs almost always involve networks of tele-immersion (the exception here is where several people share the same physical space in the VE, say, standing together in an IPT).

The networks we belong to thus extend our relations to the meso- and then to the macro-level, where populations are made up of overlapping network memberships. On this level, as several chapters in this volume document (Chapters 4, 5, 6), users experience what these authors describe as a sense of “community” in different ways. I put “community” in inverted commas since this term seems to imply strong and positive ties, whereas “networks” is more neutral and also includes the weak ties that often characterize CMC [37]. Analyzing networks is useful because it allows us to identify the boundaries of networks, and to address issues such as: who has access to particular networks, and with what kind of technological capabilities and resources? And what is the density or strength, or the weakness or diffuseness, of networks?

This brings us to stratification, and to the larger question which has often been posed in connection to CMC and shared VEs, and which Axelsson takes up in Chapter 11: whether CMC equalizes the status of participants because of the absence of social cues and other status markers? Yet, as Axelsson’s chapter shows, in shared VEs, the effect can just as often be to amplify stratification in new ways. In shared VEs stratification and hierarchy depend, for example, on the extent to which individuals can display their unique status characteristics so that they are recognizable by other participants. Again, this depends partly on familiarity with the VE: Chapters 3 and 11 give a number of examples where these characteristics are recognizable only to certain participants, such as experienced users [20].

Again, these social markers are mediated by the frame of interaction: how much attention will participants pay to status markers? How much will they trust them as reflecting the characteristics of the “real” users? And from the side of the environment, this will also depend on the geography of the VE: how much access do people have to each other? How do different spaces foster the density of networks by bringing people together in a shared VE space, or do they segregate them into different worlds and thus promote more differentiated or diffuse networks?

I have only begun to sketch a framework for analyzing interaction in shared VEs and given a few examples of how this framework applies to various findings. Much more would be needed to fill in the details of this framework, but these gaps will also have to be filled with many more empirical studies which add to our knowledge of different types of shared VEs. The study of shared VEs is still at an early stage but, as this volume shows, it will be useful to start bringing our research together in order to improve the technology and learn from its uses.

References

1. Schroeder, R. (1996). *Possible Worlds: The Social Dynamic of Virtual Reality Technology*. Boulder, CO: Westview.
2. Ellis, S. (1991). Nature and Origins of Virtual Environments: a bibliographical essay. *Computer Systems in Engineering* 2(4): 321-47.
3. Cruz-Neira, C., Sandin, D., DeFanti, T. (1993). Surround-screen projection-based virtual reality: the design and implementation of the CAVE. *Computer Graphics Proceedings, Annual Conference Series 1993*. ACM SIGGRAPH, New York: ACM, pp.135-42
4. Smith, M. and Kollock, P. (Eds) (1999). *Communities in Cyberspace*. London: Routledge.
5. Kiesler, S. (Ed) (1997). *Cultures of the Internet*. Mahwah NJ: Lawrence Erlbaum.
6. van Dijk, J. (1999). *The Network Society: Social Aspects of New Media*. London: Sage.
7. Schroeder R., Heather, N., Lee R. (1998). The Sacred and the Virtual: Religion in Multi-User Virtual Reality. *Journal of Computer Mediated Communication*, Volume 4, Issue 2, (available online).
8. Isselstein, W.A., Lombard, M., Freeman, J. (2001). Toward a Core Bibliography of Presence, *Cyberpsychology and Behaviour*, 4(2): 317-321.
9. Scheumie, M.J., van der Straaten, P., Krijn, M, van der Mast, C. (2001) Research on Presence in Virtual Reality: A Survey, *Cyberpsychology and Behaviour*, 4(2): 183-201.
10. Ellis, S. (1996). Presence of Mind: A Reaction to Thomas Sheridan's "Further Musings on the Psychophysics of Presence", *Presence: Journal of Teleoperators and Virtual Environments* 5(2): 247-59.
11. Slater, M., Sadagic, A., Usoh, M., Schroeder, R.(2000). Small Group Behaviour in a Virtual and Real Environment: A Comparative Study. *Presence: Journal of Teleoperators and Virtual Environments*, 9(1): 37-51.
12. Axelsson, A., Abelin, Å., Heldal, I., Schroeder, R., Wideström, J. (2001). Cubes in the Cube: A Comparison of a Puzzle-solving Task in a Virtual and a Real Environment. *CyberPsychology and Behavior*, 4(2): 279-86.
13. Schroeder, R. et al. (2001). Collaborating in networked immersive spaces: as good as being there together? *Computers & Graphics*, 25 (5): 781-88.

14. Axelsson, A., et al. (1999). Collaboration and Communication in Multi-User Virtual Environments: A Comparison of Desktop and Immersive Virtual Reality Systems for Molecular Visualization, in T.Fernando (ed.) *UKVRSIG'99. Proceedings of Sixth UKVRSIG Conference*, University of Salford 14th Sept., pp.107-18.
15. Schroeder, R., Axelsson, A. (2000). Trust in the Core: A Study of Long-term Users of Activeworlds, in *Digital Borderlands, a Cybercultural Symposium*, Norrköping, Sweden, 12-13 May.
16. Buescher, M., O'Brien, J., Rodden, T. and Trevor, J. (2001). "He's Behind You": The Experience of Presence in Shared Virtual Environments. In Churchill, E., Snowdon, D. and Munro, A. (Eds). *Collaborative Virtual Environments: Digital Places and Spaces for Interaction*. London: Springer, pp.77-98.
17. Allwood, J., Schroeder, R. (2000). Intercultural Communication in Virtual Environments, *Intercultural Communication*, issue 3 (available online).
18. COVEN <http://coven.lancs.ac.uk>
19. Abelin, Å., Axelsson, A., Schroeder, R. (2001). Anyone speak Spanish? Language encounters in virtual environments - a study of Active Worlds, in *The Seventh International Symposium on Social Communication*. Santiago de Cuba.
20. Schroeder, R. (1997). Networked Worlds: Social Aspects of Multi-User Virtual Reality Technology. *Sociological Research Online*, 2(4) (available online).
21. Hirose, M., Ogi, T. and Yamada, T. (1999). Integrating Live Video for Immersive Environments. *IEEE Multimedia*, July-September: 14-22.
22. Schroeder R., Huxor A., and Smith A. (2001). Activeworlds: geography and social interaction in virtual reality. *Futures: a journal of forecasting, planning and policy* 33: 569-87.
23. Dodge M. and Kitchin R. (2000). *Mapping Cyberspace*. London: Routledge.
24. Munro, A., Höök, K. and Benyon, D. (Eds) (1999). *Social Navigation of Information Space*. London: Springer.
25. Schroeder, R. (1999). Social Life in Virtual Worlds: Structure and Interaction in Multi-User Virtual Reality Technology. *Communications & Strategies*. 33 (1): 137-150.
26. Craven, C. et al. (2000). Ages of Avatar: Community Building for Inhabited Television. In E. Churchill and M. Reddy (Eds.) *CVE2000: Proceedings of the Third International Conference on Collaborative Virtual Environments*. New York: ACM Press, pp.189-94.
27. Turner, J. (1988). *A Theory of Social Interaction*. Stanford CA: Stanford University Press.
28. Collins, R. (1988). *Theoretical Sociology*. New York: Harcourt Brace Jovanovich.
29. Pargman, D. (2000) *Code Begets Community: on social and technical aspects of managing a virtual community*. Linköping studies in arts and science no. 224: Linköpings Universitet.
30. Turkle, S. (1997 2nd ed.) *Life on the Screen: Identity in the Age of the Internet*. London: Phoenix.

31. Stanney, K. and Salvendy, G. et al. (1998). Aftereffects and sense of presence in virtual environments: Formulation of a research and development agenda. *International Journal of Human-Computer Interaction*. 10: 135-87.
32. Goffman, E. (1959). *The Presentation of Self in Everyday Life*. New York: Doubleday.
33. Goffman, E. (1974). *Frame Analysis*. New York: Harper and Row.
34. Meyrowitz, J. (1985). *No Sense of Place: The Impact of Electronic Media on Social Behaviour*. Oxford: Oxford University Press.
35. Benford, J. et al. (1995). Networked Virtual Reality and Cooperative Work. *Presence: Journal of Teleoperators and Virtual Environments* 4(4): 364-86.
36. Schiano, D. (1999). Lessons from LambdaMOO: A Social, Text-Based Virtual Environment. *Presence: Journal of Teleoperators and Virtual Environments* 8(2): 127-139.
37. Wellman, B., Gulia, M. (1999). Net-Surfers Don't Ride Alone: Virtual Communities as Communities. In Wellman, B. (Ed.) *Networks in the Global Village*. Boulder, CO: Westview, pp.331-66.