Natural Object Focussed Collaboration in Distributed Virtual Environments

O. Otto

Ph.D. Thesis **2006**

Natural Object Focussed Collaboration in Distributed Virtual Environments

Oliver Otto

Centre for Virtual Environments,
Informatics Research Institute (IRIS),
University of Salford, UK

Submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy,

August 11, 2006

for my family

Preface

The University of Reading had a track record of researching Collaborative Virtual Environments when I joint it, and it had just installed a CAVE so that the impact of immersion on remote collaboration could be assessed. I was on the team that would undertake this work and my contribution is described in this thesis which I submit as partial fulfilment of my PhD. The unique nature of the CAVE lends this technology well to support human communication. The user is physically placed within the virtual environment with respect to remote users and objects. Body movement (usually the head and hands) is continuously tracked, allowing both conscious and unconscious non-verbal communication to be captured. Therefore we decided to develop an application that makes use of this technology but not only as a single user application, as done by so many others, but as a team application. Although desktop collaborative virtual environments have been well studied, few attempts to support collaboration between immersed users and in particular CAVEs, had been made. Furthermore, none had allowed remote users to share manipulation of the same object.

Halfway through the PhD we moved to the University of Salford with the opportunity to continue our research on the greatest collection of immersive displays we could hope for. At the same time we had built-up connections to a number of European CAVE research groups which we used for our research in various trials. For myself, I come with a technical background and this PhD with its sociological implications created opportunities for me to develop a deeper understanding of a large number of related topics. Some of them could have been developed into their own PhD, as I only scratched the surface. Although only an incremental step in understanding, I hope that this work will generate a continued interest in many aspects of its content.

Acknowledgements

During these years so many people helped me in so many ways, and I wish it was possible to thank them all. First and foremost I would like to thank my supervisor Dave Roberts, who was guiding me over the years to improve my style of writing and to help me focus when it was needed. In this respect I would like to thank Robin Wolff as my co-developer in this research for his support and endless patience. Furthermore, this work would not have been possible without the many people that volunteered to help in the user trials and I would like to take here the chance to thank all of them.

Secondly, I would like to thank Irina Sander for her immense support over the last year, helping me to focus on finishing this work and our endless conversations that had great influence on both the content and clarity of this work. Here I also would like to give my greatest gratitude to Farshid Amirabdollahian who now for the second time after the MSc has also read through my PhD to improve on its English as well as content.

Not to forget are all those people I have meet over the years and with whom I had great fun, making this time so much more enjoyable, in no specific these were Robin, Farshid, Diona, Lori, Rui, Silvia, Outi, Guille, Irina, Michael, Chris and many more.

Schließlich moechte ich auch meiner Familie danken, die mich ueber all die Jahre unterstuetzte und fuer die es sicher nicht einfach war das ich so lange in einem anderen Land gelebt habe. Daher moechte ich hiermit ihnen diese Arbeit widmen – This work is dedicated to my family.

Abstract

In an increasingly global economy there is rising pressure to expand collaboration from co-located to geographically distributed groups. Currently natural human interaction is not well supported between people collaborating across a distance. This negatively impacts on the feeling and performance of collaboration. Cooperative working could be better supported by richer mediums with more natural interfaces that allow people to interact with shared objects and each other as if they were co-located. For example, intention and opinion must be communicated, while synchronously manipulating shared artefacts. Transferring the straightforwardness of such collaboration onto distributed teams is challenging. Various forms of teleconferencing systems attempt to offer such support, yet they have difficulties with sharing objects and the direct social response this involves when participants interact with those objects. This work demonstrates that a collaborative virtual environment (CVE) can assist such cooperation and that immersive displays are of greater help compared with the traditional desktop interfaces to bring us closer to replicating a face-to-face interaction. The effectiveness of application of this technology depends on a complex set of factors that determine the efficiency of collaboration. This work examines these factors and their interrelationships within the framework of a taxonomy focussed on supporting closely-coupled collaboration using immersive CVEs.

Contents

List of Figures

Number *Page*

List of Tables

Number *Page*

Published contributions

Journal Papers:

Roberts, D., Wolff, R., and Otto, O. (2005). Supporting a Closely Coupled Task between a Distributed Team: Using Immersive Virtual Reality Technology. *Computing and Informatics,* 24, pp.1001-1023.

Wolff, R., Roberts, D., & Otto, O. (2004). A study of event traffic during the shared manipulation of objects within collaborative virtual environments. *Presence: Teleoperators & Virtual Environments,* 13(3), pp.251-262.

Roberts, D. J., Wolff, R., & Otto, O. (2003). Constructing a Gazebo: Supporting team work in a tightly coupled, distributed task in virtual reality. *Presence: Teleoperators & Virtual Environments 2003,* 12(6), pp.644-657.

Book Chapter:

Roberts, D., Wolff, R., and Otto, O. (2005). *The impact of display system and embodiment on closely coupled collaboration between remote users.* In Schroeder, R., Axelsson, A.S. (ed.) Avatars at Work and Play: Collaboration and Interaction in Shared Virtual Environments, 2005, London, Springer-Verlag, pp.131-150.

Papers:

Otto, O., Roberts, D., and Wolff, R.(2006), *A Review on Effective Closely-Coupled Collaboration using Immersive CVEs*, In Proceedings of ACM SIGGRAPH International Conference on Virtual Reality Continuum and Its Applications (VRCIA), CUHK, Hong Kong, 14-17 June, pp. 145-154

Otto, O., Roberts, D., and Wolff, R.(2005), *A Study of Influential Factors on Effective Closely-Coupled Collaboration based on Single User Perceptions*, In Proceedings of the 8th Annual International Workshop on Presence, London, UK, pp. 181-188

Roberts, D., Al-Liabi, M., Wolff, R., Otto, O. and Al-Khalifah, A.(2005). *Reducing Fragmentation in Telecollaboration by Using IPT*, In Proceedings of IPT & EGVE Workshop, pp.211-216

Otto, O., Roberts, D., and Wolff, R.(2005), *Analysing Display Influence on a Close Coupled Task based on a Single User Evaluation*, Presented at the 2nd International Telework Academy Conference, Liverpool, UK

Roberts, D., Wolff, R., Otto, O., Kranzlmueller, D., Anthes, C., and Steed, A.(2004, November 10-12). *Supporting Social Human Communication between Distributed Walk-in Displays*, In Proceedings of VRST'04, Hong Kong, pp. 81-88.

Wolff, R., Roberts, D., & Otto, O. (2004, October 21-23). *Collaboration around Shared Objects in Immersive Virtual Environments*, In Proceedings of The Eighth IEEE International Symposium on Distributed Simulation and Real Time Applications, DS-RT 2004, Budapest, Hungary, pp. 206-209.

Seelig, M., Roberts, D., Harwin, W., Otto, O., & Wolff, R. (2004, September 15-17). *A haptic interface for linked immersive and desktop displays: Maintaining sufficient frame rate for haptic rendering.* In Proceedings of the 17th International Conference on Parallel and Distributed Computing Systems (PDCS 2004), San Francisco, CA, USA, pp. 478-483.

Roberts, D., Otto, O., & Wolff, R. (2004, September 15-17). *Pushmepullyou: The reality of interaction with shared objects in networked walk-in displays.* In Proceedings of the 17th International Conference on Parallel and Distributed Computing Systems (PDCS 2004), San Francisco, CA, USA, pp. 470-477.

Otto, O., & Roberts, D. (2003, October 23-25). *Importance of Communication Influences on a Highly Collaborative Task.* In Proceedings of the The Seventh IEEE International Symposium on Distributed Simulation and Real Time Applications, DS-RT 2003, Delft, The Netherlands, pp. 195-201

Poster:

Otto, O. (2004, April 21). *Social human communication during shared object manipulation* Poster presented at the CVE IRIS open day, The University of Salford

Chapter 1

1 Introduction

From the early phase of Man's development, communication was an important tool for success. It enabled us to share knowledge, for example, how to use tools, find resources or solve conflicts. As providing food through teamwork became easier, it also allowed for more time to socialise. Communities developed which eventually evolved into civilisations. One day, knowledge was no longer spread through demonstration and speech alone, but also by writing. Later the development of the book print accelerated the knowledge transfer. Subsequently it led to the industrial revolution and people were even more required to work together as production processes became too complex to be handled by one person alone. Today teamwork is a requirement for nearly every job and through globalisation it becomes more and more distributed.

Data sharing is a common practice in many disciplines, such as design, science, education, finance, and medicine. Due to the broaden nature of resources and expertise, there has to be a need for such data to be shared amongst distributed users located in different geographical locations. In order to exchange experiences, get feedback, offer assistance, and "give-a-hand" to others, there is a need to find ways to share such representations. Medicine is a typical example of such scenarios. Participants may need to share models or real patient data. At the professional level medical experts may need to share data for diagnostic or counselling purposes. Similarly on the academic level, sharing models can be a valuable procedure in demonstrating anatomy and structures.

In the real world, people perceive and interact with their environments via what is called the "Social Human Communication" (SHC) paradigm, discussed in detail in Chapter 2. In brief SHC encompasses a dichotomy of interactions composed of a number of communication forms. These are verbal, non-verbal, and the role of objects and environment on communication. Communication is not always abstract, but normally relates to people's surroundings and artefacts, which both provide a cue for understanding. People may discuss their surrounding or an artefact via verbal and non-verbal communication, but they also like to move around their environment and interact with its contents.

Data visualisation is more adequate when users are able to interact with it, to allow enriched data exploration. Moreover, it is desirable if users can collaborate with and around this data. Thus, there is a need to communicate at a human level while sharing such data. The role of this communication is far more imperative when interactions are the core elements of the system. In order to understand complex data and take full benefit of the technology involved, there has to be a mechanism for proper human communication and interaction, especially for distributed users. Different technologies try to support such collaboration and this thesis studies the effects of immersive displays on closely-coupled collaboration and social human communication.

1.1 The importance of remote collaboration

In a globalising world large corporations are becoming increasingly distributed for a number of possible reasons and broadly speaking this could be split into five motives. First, mergers and acquisitions to adjust and complement product lines often lead to new sites becoming part of the company. Second, to participate in some markets government regulations request the location of some local development operations. Third, it can make sense for market reasons to locate parts of the corporation where the market for a particular technology exists. Fourth, the competition for highly skilled technical staff is driving companies to hire them wherever in the world the talent can be found. Finally, most corporations, especially those in the software business, hope that geographic distribution could lead to round-the clock development, which offers the promise of reducing development cycles by increasing the amount of daily development. For example, the working day in Australia does not overlap with the working day in the United Kingdom, making it theoretically possible to get 16 or more hours of development in one day.

However, not only large corporations are increasingly distributed, but also cooperating research groups and other organisations. For most tasks the communication and cooperation of these distributed groups can be achieved by using uni-modal technology (telephone, email, message boards). But a number of tasks require a higher level of collaboration including synchronous interaction in order to gain better trust, understanding and to faster resolve different views. Normally this can be achieved via face-to-face interaction, but for distributed groups this is not always an option and technology is used to resolve this problem. Ellis et al. classified interaction according to whether it is synchronous and co-located (see Table 1-1) in a ‘time space’ groupware taxonomy [Ellis et al., 1991]. A detailed description and discussion of those technologies can be found in Chapter 3.

Table 1-1: *Ellis et al. 'time space' groupware taxonomy, classifying interaction according to whether it is synchronous and co-located*

	same time	**different times**
same place	face-to-face interaction (e.g. meeting room)	asynchronous interaction (e.g. physical bulletin board)
different places	synchronous distributed interaction (e.g. video-conferencing)	asynchronous distributed interaction (e.g. Email)

In the field of computer supported distributed work (CSCW), a great deal of confusion is caused by the different interpretations of the terms collaboration and cooperation as many authors simply consider both terms as synonyms, while others draw a distinction between them [Dillenbourg et al., 1995]:

Cooperation and collaboration do not differ in terms of whether or not the task is distributed, but by virtue of the way in which it is divided; in cooperation the task is split (hierarchically) into independent subtasks; in collaboration cognitive processes may be (heterarchically) divided into intertwined layers. In cooperation, coordination is only required when assembling partial results, while collaboration is « ...a coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem ».

Following the analogy of a globalising world, the cooperation between groups is usually a common task. Some of these tasks involve the designing of objects, simulating of environments, training collaborating teams and visualising data. A common factor between all these four categories is their need to the visualisation of 3 dimensional (3D) data which can be achieved by using the Virtual Reality (VR) technology.

1.2 Reasons to use Virtual Environments for Collaboration

Visualisation technologies, such as virtual environments (VE), allow scientists and academics to investigate models and data in a more intuitive manner. It transfers the data into graphical representations, which can be easily interpreted and understood. Visual inspection allows identifying patterns and irregularities. The need for VEs is multifaceted, for example, if you cannot get/realise something in a real environment, because it:

- is not here (local) or does not exist at all,
- is too expensive to buy/rent/create,
- is too dangerous to perform,
- takes too much effort to get/realise/perform,

- takes too much time to get/realise,
- is currently not possible to create or to get into it
- is not visible, because too small/big/wrong frequency spectrum
- is too fast/slow or no longer/not yet there (period/age)

These are a few reasons to use a virtual environment to ‘fake reality’ by simulating (visual/aural/touch) it. These environments could be used in a single user mode or more productively shared by many and used in collaboration. Most visualisation applications currently only support asynchronous collaboration (see groupware in Chapter 3), but a few technologies are designed for concurrent collaboration which includes Collaborative Virtual Environments (CVE). For example, CVEs are used when more than one user is needed to perform a task, to explore or discuss things in the virtual world. For collaborating teams it is not only important to visualise their problem, task, subject or idea but also to synchronously interact with it.

1.3 Closely-coupled Collaboration

Collaboration between people is often centred around their common interests. These points of interests may be embodied by some perceivable objects. If this collaboration requires verbal and non-verbal communication including shared concurrent object manipulation, then we class it as closely-coupled collaboration. To be precise, closely-coupled collaboration is a close coupling between object manipulation and human interaction, whereas the action of collaborating people is directly depending on each other. Shared manipulation by multiple users can be sequential or concurrent of the same and through different attributes of objects as well as instructions to others (Table 1-2). It is important for all collaborators to perceive and understand the object in order to work with it. While we cooperate with other people through an object, we use a variety of communicational resources to demonstrate our opinion, intention and needs to others. Be it simply verbally with emotional nuances, with gestures and postures in a non-verbal way or by manipulating the object directly. When interacting remotely, these forms of social human communication (SHC), as well as the representation of the object, need to be mediated through tele-collaboration technology.

Table 1-2: *Forms of shared object manipulation with respect to timing*

Timing	Method of sharing	Example
asynchronous	sequential manipulation of **distinct** object attributes	a person moves an object to a place, then another person paints it
	sequential manipulation of the **same** object attributes	a person moves an object to a place, then another person moves it further
synchronous	concurrent manipulation of **distinct** object attributes	a person is holding an object while another person is painting it
	concurrent manipulation of the **same** object attributes	several people lift a heavy object together

1.4 Motivation, Challenges and Research Objectives

Currently natural human interaction through and around objects is not well supported between people collaborating across a distance and a motivation of this work was to improve our understanding of how this can be achieved. Many teamwork tasks require a close coupling between the interactions of members of a team. For example, intention and opinion must be communicated, while synchronously manipulating shared artefacts. In face-to-face interaction this communication and manipulation is seamless. Transferring the straightforwardness of such collaboration onto remote located teams is technologically challenging. To implement closely-coupled collaboration for remote collaborators the usage of CVEs is most promising. It has the potential to overcome issues encountered when using other technologies (see Chapter 3) and is most suitable for remote interaction and visualisation of arbitrary environments. CVEs enable people in remote locations to interact with synthetic objects and representations of other participants within. A study by Broll in the mid-nineties, concluded that concurrent shared manipulation of objects in a CVE would not be possible with technology at that time, due to delays caused by distribution [Broll, 1995]. In following years, a number of studies were able to demonstrate such interaction. At first using single desktop system [Ruddle et al., 2002] and later through networked immersive displays [Mortensen et al., 2002; Linebarger et al., 2003; Roberts et al., 2003]. Demonstrating closely-coupled collaboration is an important step towards an application allowing for many human interactions. Supporting such interactions would allow a wider range of activities across a distance in applications as diverse as surgical planning to design review meetings or emergency simulations. Before such technology can be realised, application developers and system designers need to understand how to support effective closely-coupled collaboration. The answer is not easy and depends on a variety of factors (Figure 1-1), a number of which are discussed in this thesis. This work will focus on immersive systems as we believe they hold the key for effective distributed collaboration.

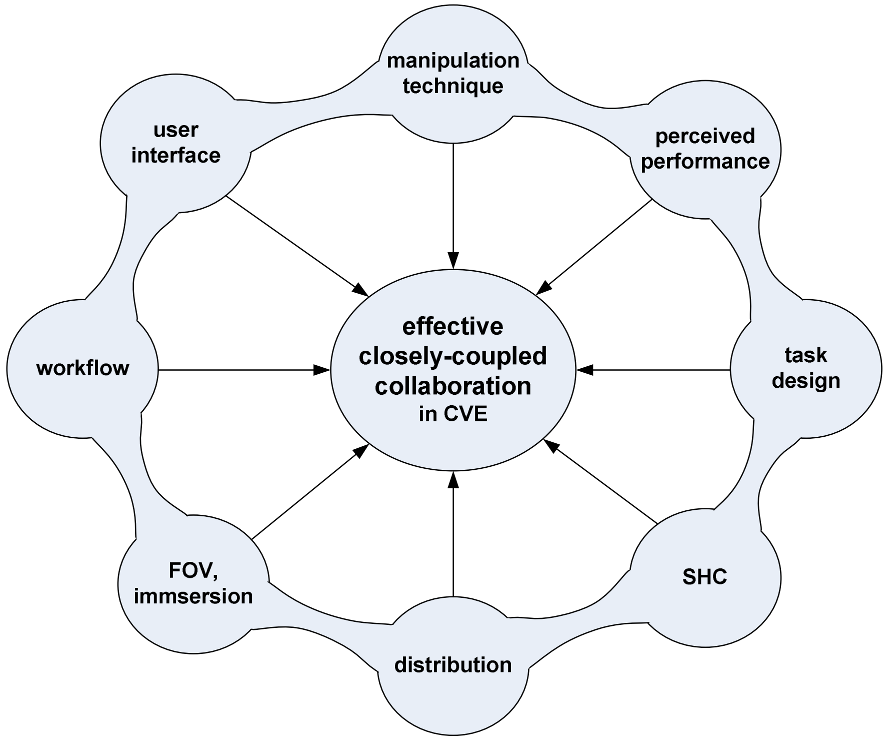


Figure 1-1: *Influences on effective closely-coupled collaboration, the different factors influence not only the effectiveness but also each other, some factors are well documented others less, but few in the context of supporting closely-coupled collaboration*

1.5 Delimitations

The research in this thesis is focused on social and technological aspects (display properties) of closely-coupled collaboration using CVEs. Various display combinations were trialled using a self-developed benchmark application. However, the focus of this thesis is not the impact on the underlying software systems while supporting closely-coupled collaboration as this was the research focus of other parties within this research group. In contrast, studies presented in this thesis based on experiments using desktop systems as well as semi- and fully-immersive displays (CAVE, Workbench, HMD) that employ a motion tracking system. The impact of haptics was not part of this study nor was the intention to perform a deep psychological analysis.

1.6 Methodology

The research for this thesis was conducted as part of a research group investigation, were each member had different focuses such as social aspects and technical issues supporting distributed closely-coupled collaboration. An ethnographic analysis has been used to gather data [Bowers et al., 1996; Crabtree, 2003], for example, measuring user behaviour, statistical analysis of questionnaires, and user observation. A number of hypotheses (Chapter 5-7) where formulated and put to the test using a self-developed benchmark application (Chapter 4). The benchmark was continuously developed to incorporate gained knowledge and to focus on further research questions (hypotheses). Research trials were conducted either locally or to collaborating partners in other Universities, involving mostly student volunteers. Some related studies by the same research group are considered to further improve and strengthen the emerging framework.

1.7 Contribution to the field

This work contributes to the knowledge with an improved understanding of the potential of immersive displays as a tool for collaboration and human interaction. It demonstrates that closely-coupled collaboration is possible and will further show that immersive CVEs can support a seamless flow of collaboration and communication. In addition, interrelationships between a variety of factors are discussed throughout this thesis, which are summarised in a framework of influences on closely-coupled collaboration.

1.8 Organisation of this Thesis

The structure of this thesis is methodical by introducing first a number of aspects in the light of closely-coupled object interaction. This is necessary to gain a better understanding of the complex issues that relate to this research. Its climax can be found in the last chapter were all related aspects are looked at again and summarised in a new framework. The work presented in this thesis is basic research building on top of long studied aspects but looking from a new angle of allowing a distributed team to interact through objects with similar body movements to which they would use in the real world, thus making it easier to communicate intention, action and emotion.

This thesis is organised as follows:

- Chapter 1 provides a quick overview of the research discussed in this thesis and presents a framework of structure for the discussion.
- Chapter 2 introduces the notion of presence and social human communication and their relevance on closely-coupled collaboration. These chapters are later used as background information for discussion and evaluation.
- Chapter 3 evaluates various technologies for remote collaboration and explains why immersive displays are of greater help than traditional desktop interfaces or other tele-working technologies to bring us closer to replicating distributed face-to-face interaction.
- Chapter 4 explains the choice of CVE and demonstrates the evolution of the benchmark application (The Virtual Gazebo). Successes and failures of this benchmark are discussed and evaluated.
- Chapter 5 discusses the first user evaluation of the benchmark application during a CAVE-Desktop-Desktop trial and the first-hypothesis that immersive CAVE-like displays are suited for closely-coupled collaboration.
- Chapter 6 looks into the evaluation of a CAVE-CAVE trial and testing of the second-hypothesis that performance increase and easier human interaction can be achieved with purely immersive interaction.
- Chapter 7 investigates how various factors influence interaction by focussing on a single user trial on three different displays (desktop, workbench, CAVE). A modified version of the benchmark application is used and various display properties determine effectiveness of collaboration (third hypothesis) are discussed. Furthermore, a related study comparing CAVE-HMD is discussed.
- Chapter 8 is summarising the various influencing factors within a framework that determines the usefulness of distributed human interaction using immersive CVEs.
- Chapter 9 concludes this work and illustrates the possibilities for future work.

1.9 Summary

This introductory chapter has introduced the notion of closely-couple collaboration and its relevance to human interaction. It thereby has provided the contextual and structural framework for the remainder of the thesis. Contextually, remote collaboration is of utmost importance for most industries and providing a tool that allows natural communication and collaboration through and around objects can pave the way for a variety of application. In addition, immersive displays to this day are very effective in placing a user into a different world and with appropriate tools allow communication and interaction around artefacts. Structurally, the roadmap leads from foundational social human issues to technological challenges and later to possible solutions and implications. It is especially the foundational issues that will be addressed in the following chapters.

Chapter 2

2 Presence and Social Human Communication

This chapter explores presence and social human communication, and their influence onto closely-coupled collaboration. This background work is used throughout the thesis as a medium for discussion and to demonstrate technological influences onto human collaboration.

2.1 Presence and Tele-presence

Awareness of oneself and others is a defining character of face-to-face collaboration. This awareness is partially created through a feeling of presence and co-presence. Therefore the recreation of this perception through technology is likely to be important for successful distributed collaboration. First, however, an understanding of the term presence and its creation through immersion has to be developed. The following section discusses types of presence, its measures and relationships to performance and collaboration. In later chapters this knowledge is used to understand and evaluate the influence of technology on distributed closely-coupled collaboration.

2.1.1 Presence and Immersion

When people talk about presence they sometimes also call it immersion, but here a distinction is made. The term "immersion" describes the extent to which a given technology replaces real world stimuli with synthetic stimuli within the virtuality continuum [Milgram & Fumio, 1994]. A necessary condition is Ellis' notion of a Virtual Environment (VE) as a communication media [Ellis, 1996], maintained in at least one sensory modality (typically the visual). The degree of immersion is increased by increasing the field of view (FOV), greater degree of body tracking, decreased lag between body movements and resulting changes in sensory data amongst others [Pausch et al., 1997; Sheridan, 2000; Baños et al., 2004]. Immersion may lead to a sense of presence and some consider it a precondition. This is a psychological emergent property of an immersive system, and refers to the participant's sense of "being there" in the world created by the VE system [Slater et al., 1994]. Immersion is a necessary rather than a sufficient condition for presence - immersion describes a characteristic of the technology, whereas presence describes an associated state of consciousness [Slater & Steed, 2000b].

The definition of presence is usually the sense of “being in an environment other than the one you are physically in”, but there is no commonly agreed theory of presence nor are there common measures for this construct. Some evidence points in the direction that particular technological, task and personal characteristics can influence the extent of presence experienced in a VE. A small set of studies have attempted to examine the relationship between presence and task performance and some show that there is a significant correlation between both. As yet it is unclear whether the relationships are causal in nature.

It is not within the scope of this section to survey all factors that may influence the experience of presence, such can be found in a very informative survey by [Draper et al., 1998], [IJsselsteijn et al., 2001] and [Youngblut, 2003]. It is rather in the interest of this section to briefly inform about the current status and complexity of presence research and its relevance to closely-coupled collaboration, task performance and dependency of technological factors such as display, immersion or interface.

2.1.2 Types of presence

The construct presence exists in four different adaptations, which are place-presence, co-presence, social-presence and object-presence. Place-presence is what VR papers normally call simply “presence” whereas all other types of presence are called by their full name.

(place) presence: This is the subjective experience of being in one place or environment, even when one is physically situated in another place or environment and it is summarised as a feeling of “being there”. One can feel presence when reading a book, watching a movie or by immersing into a virtual environment.

co-presence: This is the subjective experience of being together with others in a virtual environment, even when participants are physically situated in different sites. This is also referred to as “being there with others”.

social-presence: This occurs when users feel that a form, behaviour, or sensory experience indicates the presence of another individual. The amount of social presence is the degree to which a user feels access to the intelligence, intentions, and sensory impressions of another [Biocca, 1997].

object-presence: This is the subjective experience that a particular object exists in a user's environment, even when that object does not [Stevens et al., 2002].

2.1.3 Measure of presence

The measurement of presence is difficult as there is no commonly agreed theory nor are there common measures for this construct. An ideal measure should be:

- non-intrusive, that is, doesn't interfere with the task being performed or measurement itself
- free from a participants or experimenter bias
- easy to use, not imposing an unwarranted burden in terms of time and/or special equipment
- capable of measuring temporal variations in the construct being measured
- recording of real-time changes in the sense of presence

As the term presence is subjective and difficult to measure, various measurements have been established over the past decade. The typical approaches are presence questionnaires (e.g [Singer & Witmer, 1999; Slater et al., 2000b]). They use questions which are given to the participants of a user trial and focus on various aspects of presence. However, as participants have no real common understanding of the term "presence", many related questions have to be asked to reduce errors and to find an overall relationship. In addition, the answers are subjective self-assessments and a substantial large number of participants should be questioned in order to find an average value with low entropy.

Another way to measure presence is based on discriminating between environments [Welch et al., 1996; Snow & Williges, 1998; Slater et al., 2000b], in which the major component of the measure depends on data collected during the course of the VE experience itself. For example, Slater et al. based their measure on the number of transitions between the state of being in the VE to the state of being in the real world.

Sometimes psychological measurements [Darken et al., 1999; Kalawsky, 2000; Nichols et al., 2000; Sas & O'Hare, 2003] are used to decompose presence into measurable subcomponents by looking into the complex interactions of the human sensory and perceptual systems with a stimulus environment. A more objective approach are physiological measurements [Jorgensen et al., 1997; Wiederhold et al., 1998; Meehan et al., 2001] where participants' body reactions are measured (e.g. heart rate, respiration rate, peripheral skin temperature). Other measurements that also do not rely on subjective interpretations are observed reactions [Slater & Usoh, 1993; Tromp et al., 1998] to a given stimuli.

Finally, after conducting an experiment post-interaction effects are measured to determine the influence of technology and presence on motion sickness-like symptoms and other aftereffects such as disorientation, visual stress or altered hand-eye coordination [Welch, 1997; Stanney & Salvendy, 1998]. As the technology develops, and the implementation of VE systems migrates to widely available platforms (e.g., the internet), understanding the long-term ramifications of aftereffects and presence in terms of ergonomic applications will become ever more important. While it is undeniable that some people experience some aftereffects within some systems, it is not necessarily the case that VE technology is implicitly harmful.

The measurement of co-presence is desirable when interacting and sharing an environment with others. Although all measures described above could be used to study effects on co-presence, the effort to do so is higher for some compared to others. Therefore mainly co-presence questionnaire [Axelsson et al., 1999; Slater et al., 2000a; Axelsson et al., 2001; Casanueva, 2001] are used in the majority of studies. The lack of research in this area is surprising, giving the expected wide use of collaborative VEs. One of the problems is that current presence measures are intended to assess an individual's presence experience. This requires identifying performance measures that relate to individual team performance in a meaningful manner. Another problem is that a larger number of subjects are needed to generate statistically significant results. When team-oriented performance measures are used, not only more participants are needed to form the teams, but the inevitable variation among participants across teams is difficult to control. Just as with co-presence the study of social-presence is rather limited and again the method of measurement are questionnaires [Thie & Wijk, 1998; Sallnäs et al., 2000; Bailenson et al., 2001].

Questionnaires and observations are used during this research to measure the influence of presence on closely-coupled collaboration. Later Chapters will demonstrate that task design and display setup (see Chapter 5&6) as well as the type of display have an influence (see Chapter 7).

2.1.4 Relationships of presence

As is known the construct presence is difficult to define and it is just as difficult to determine what influences the feeling of "presence". There is, however, some evidence that particular technological, task and personal characteristics can influence the extent of presence experienced in a VE. In addition the various types of presence can have interrelationships. For example many agree that place-presence is a precondition for co-presence [Slater et al., 2000a] or if participants are together with others one can collaborate and again the argument is that co-presence is a precondition for collaboration [Tromp et al., 1998].

Only one study made the effort to look into the issue of social presence and co-presence and found a significant correlation [Thie et al., 1998]. The same study states also that social presence is a sub-mental model of (place)-presence.

2.1.4.1 Synopsis of technological characteristics and presence

It is not the scope of this section to survey all factors that may influence the experience of presence, for this a survey by Youngblut is very informative [Youngblut, 2003]. However, the number of factors that may influence the experience of presence is large and at least 30 factors can be distinguished. This includes avatar realism, display resolution, field of view (FOV), haptic cues, interaction (level of), scene realism, social presence, tactile cues and many more. The overall findings are inconsistent as are the methods of measurement. However, consistently significant characteristics were found to be FOV, 3D vision, update rate, dynamic shadows and frame rate. The majority of findings were positive for a significant relationship with head tracking, navigation and participant movement. For other characteristics the findings were inconsistent and further examination ought to be done [Youngblut, 2003].

2.1.4.2 Synopsis of technological characteristics and co-presence

Youngblood's survey also looked for studies investigating co-presence and found some factors that may influence this experience. These include avatar functionality, avatar realism, collaboration, environment type, haptic force feedback, presence manipulation and visual display. The overall findings are inconsistent and in most studies no significant effect was found for visual display unless both participants were immersed using a CAVE-like display [Schroeder et al., 2001].

2.1.4.3 Task characteristics

Task characteristics, in the sense used here, do not relate to the *activities* performed in a VE. Instead, they affect the *manner* in which a participant is asked to complete a task. For example, (a) whether collaboration with a partner is required or not, (b) which type of instructional technique is employed, or (c) whether a participant has multiple exposures to a VE. At first glance, some of these characteristics seem to be the same as those identified as technological characteristics. There is a difference, however, where for example task related information is enhanced with audio cues (e.g. drilling sound while using a drill tool) as opposed to the impact of constant background sounds. Other studied factors include collaboration, distance cues, elapsed time to testing, practice with interface, task complexity, task expertise and others. The reader is encouraged for more detailed studies of these factors to refer to very informative past research presented by [Draper et al., 1998; IJsselsteijn et al., 2001; Youngblut, 2003].

2.1.4.4 Relationship between task performance and presence

There are number of studies to consider and there are a total of fifty findings to review. Half of these showed significant correlations between task performance measures and presence, and over 90% of these correlations were in the expected direction that they influence task performance measures and presence [Singer et al., 1995; Snow, 1996; Welch, 1999; Nichols et al., 2000; Zimmons & Panter, 2003]. The main focuses of these studies were on accuracy, collisions, errors made and time to complete the task.

Virtual environments are a tool and can help to improve collaboration over distributed sites if used sensibly and applied correctly. It is the research communities' role to establish how to improve the effectiveness and quality of VE experience for subjects and user organizations, and to establish and justify clear performance benefits for evaluated applications.

2.1.4.5 Relationship to collaboration

Presence in VR is often used as a benchmark to see how engaging a task was and some results indicated that collaboration does not increase the sense of presence in the virtual environment, but does improve the quality of the experience in the virtual environment [Bystrom & Barfield, 1999]. However, a recent study found that both the communication media used and the environment in which collaboration takes place

(CVE or Web) make a difference on how subjects experience interaction and on their communication behaviour. Participants rated presence higher in the collaborative video condition compared to the audio only condition [Sallnäs, 2005]. Results in Chapter 7 show that presence is depending on display properties such as field of view, size and navigational freedom (see Chapter 7.6.1).

Many agree that (place-) presence is a precondition for co-presence [Slater et al., 2000a] or if participants are together with others one can collaborate and again the argument is that co-presence is a precondition for collaboration [Tromp et al., 1998]. Although some found a significant correlation between collaboration and co-presence if IPTs are used [Axelsson et al., 1999; Axelsson et al., 2001; Schuemie et al., 2001], no such significant correlation could be found for using desktops [Tromp et al., 1998; Bystrom et al., 1999; Casanueva, 2001].

When immersed pairs in CAVE-like displays were asked to carry concurrently a "stretcher", the study suggested that co-presence was significantly and positively correlated with task performance [Mortensen et al., 2002]. However, Chapter 6 (see 6.6.4) will show that co-presence is rated very high by all participants for a closely-coupled task independent of their performance. Other research found that there was a large difference in the co-presence scores between VEs of low and high collaboration. In the high collaborative scenario users had to negotiate the release of objects by standing close together, while at the same time each user was trying to finish their own task of moving objects to different rooms. The study indicated that participants in the high-collaboration VE had a much larger sense of co-presence than participant in the low-collaboration VE [Casanueva & Blake, 2000].

Current IPTs usually create reasonable audio-visual feedback, yet during interaction with VE's an important human sense is missing. The reproduction of the feeling of touch is attempted via haptic interfaces and results [Massie, 1993; Van der Linde et al., 2002; Kim et al., 2004] show that haptic force feedback significantly improves task performance, perceived task performance, and perceived virtual presence in a CVE. The results suggest that haptic force feedback increases perceived social presence, but the difference is not significant [Sallnäs et al., 2000]. Other research indicate similar results, where task performance was significantly faster and more precise when haptic force feedback was provided [Hurmuzlu et al., 1998]. In addition, haptic feedback increased the feeling of togetherness and improved task performance when pairs of people moved a ring on a wire collaboratively [Basdogan et al., 2000]. The ring only moved when both participants pushed in the same xyz-direction, similar to the task of carrying a beam in the Virtual Gazebo benchmark (see Chapter 4).

Some of the results mentioned above seem to be in contradiction. However, it appears that with the intensity of collaboration a higher perception of co-presence can be found and that presence depends on a number of properties such as technological and task characteristics. Furthermore, improved task performance and the feeling of togetherness can benefit from the support of haptics.

2.2 Importance of Social Human Communication and Trust for Collaboration

Collaboration between humans depends on both complex social interaction and the physical medium in which these interactions take place. An understanding of the nature of interaction in the real world can help to reason about collaboration in virtual environments and may lead to a definition of its requirements. Social Human Communication (SHC) encompasses a taxonomy of interaction that includes verbal and non-verbal communication, and the role of objects and the environment on communication [Knapp, 1978; Burgoon et al., 1994]. Verbal communication includes mainly the human linguistic system and its derivatives (speech, writing, sign language, etc.) and our body language, including posture and gesture, is considered as non-verbal communication. They both provide a context for understanding. Verbal and non-verbal communication are often inextricably linked through nuances such as lip-synch, clapping and unintentional gesturing and posture changes while speaking. The subject of communication is not always abstract and often relates to our surroundings and artefacts within it. People may discuss their surroundings or an object through both communication modes, but in addition they can move around the environment and manipulate objects within it. A nuance might arise from the synchronisation of concurrent elements of SHC. For example, a user might point to an object saying “lets pick that up” and then turn and point to a place in the environments saying “and take it over there”, thus relating verbal and non-verbal communication in relation to an object and the environment.

The purpose for this review is to highlight the importance of human communication in human collaboration. The implementation of a mechanism allowing users to employ SHC during their collaboration in CVEs is important for the quality of the interaction and the immersive aspect of the system. It is argued that “we can learn more about the nature of human communication by observing how it is affected by technology and correspondingly, we can learn more about the social nature of communication technologies by thinking about how they both rely upon and transform basic human communication patterns" [Hutchby, 2001].

2.2.1 Verbal communication

Techniques for analysing human-human interaction may fruitfully be applied to human-machine-human interaction. Conversation analysis (CA) is the study of talk-in-interaction [Hutchby & Wooffitt, 1998]. This is the systematic analysis of the kinds of talk produced in everyday, naturally occurring situations of social interaction. In research, a conversation is usually taped and later analysed. The majority of CA work, however, has focused audio analysis via telephone conversation, but this method can also be adapted to CA in virtual reality since interaction in virtual environments including audio can also be recorded and analysed just like a telephone conversation. In addition, video and VE moments can be recorded for behavioural analysis which then could be linked with the CA results.

We demonstrated in an example application (see Chapter 5), called the Virtual Gazebo, that speech between participants is vital. The lack of certain cues (gestures, touch, etc.) affords more verbal communication to coordinate and fulfil a task while collaboration without speech can be very difficult. For example of a simple dialog may demonstrate this problem:

Bob: Give me the plank, please. I want to fix it on the beam.
Jeff: Which plank, the green or yellow one?
Bob: Just give me one!
Jeff: Shall I place it for you while you fit it?
Bob: Yes, please!

Without speech this dialog would take time and it would be probably more effective for Bob to get the plank himself. The advantage of VR is that it can provide participants with more cues which again help them to understand and to interact (among other things it increases the feeling of co-presence). Nonetheless, verbal communication plays a vital role in closely-coupled collaboration and CA can help researchers to better understand the way in how applications have to be improved. Conversation analysis on talk-in-interaction is helping researchers to understand how people change the subject of a conversation (as used in Chapter 6), known as *turn taking*. The analysis for turn taking conversations is focused on answering the question "What and how do we change the focus of a conversation". The argument is to show that mutual understanding or intersubjectivity are publicly acceptable accomplishments, which are observable in the data of *talk-in-interaction*. Researchers have observed that interactions tend to fall into certain patterns. For example, people have developed systematic, recursive ways of beginning (and ending) telephone conversations that may exhibit differences from their ways of beginning conversations in face-to-face circumstances.

Hutchby argues that all technological objects (artefacts) have their own affordances and therefore we are frequently forced to make changes to the international conventions for conversations [Hutchby, 2001]. In other words, humans are immensely adaptive and creative in using natural conversation resources, e.g. using smiley in a text conversation. This thesis asks the question "What affordances do CVEs have and how is this affected by technological factors?".

2.2.2 Non-verbal communication

Non-verbal communication can play a supportive or even predominant role in completing a task. A major part of face-to-face interaction is supported by non-verbal communication. Gestures, bodily orientation, eye gaze and so on are a vital part of our everyday interaction with others. Goleman argues that people are not usually aware of their non-verbal behaviour [Goleman, 1996]. This display of non-verbal cues becomes a force of habit and typically occurs below the level of awareness [Ekman & Friesan, 1969]. The following is a relatively simple classification [Darn, 2005]:

Table 2-1: *Simple classification of non-verbal cues*

Kinesics	body motions (blushes, shrugs, eye movement, foot-tapping, drumming fingers)
Proxemics	spatial separation (in relation both the social and physical environment)
Haptics	touch
Oculesics	eye contact
Chronemics	use of time, waiting, pausing
Olfactics	smell
Vocalics	tone of voice, timbre, volume, speed
Sound Symbols	grunting, mmm, er, ah, uh-huh, mumbling,
Silence	absence of sound (muteness, stillness, secrecy)
Adornment	clothing, jewellery, hairstyle
Posture	position of the body (characteristic or assumed)
Locomotion	walking, running, staggering, limping
Expression	frowns, grimaces, smirks, smiles, pouting

An array of circumstances for interaction can be found, which can be differentiated on the basis of their degree of "cuelessness" [Hutchby, 2001]. For example, in face-to-face interaction participants have the fullest range of cues, while on the telephone they have the least range (only the pitch of the voice). The degree of cuelessness in CVEs falls somewhere in between face-to-face and telephone interactions. This raises questions on how technology factors impact on the level of cuelessness. This and other work has started to answer this question. For example, an application with no gesture support is less informative than one supporting gestures [Salem, 2005]. Further, it can be rationalised that immersive projection technology (IPT), such as HMDs, CAVEs, etc. (see Chapter 3), do not rely on predefined gestures, but instead use their motion tracking system to increase the number of gestures and therefore to decrease the degree of cuelessness. The data from hand and head tracker are matched to the embodiment, which allows other participants to see where users are looking and with which object they might interact. The hand tracking also allows users some simple gestures (pointing, waving, etc.). By using a tracker for each hand the variety and flexibility is increasing. The Virtual Gazebo benchmark (see Chapter 5 and 6) shows that even simple gestures are helpful. Things like directing someone to a specific tool or place are easier with the ability to explicitly point out the direction. A VR system includes usually only a hand and head tracker and further research is needed to investigate the influence of other cues like facial expressions or eye-gaze.

Hutchby argues that the concept of cuelessness is ultimately misleading because there is no evidence that telephone conversations (interaction) are less effective than face-to-face conversations and it has less to do with the coordination of turn-taking and more to do with “psychological distance” [Hutchby, 2001]. This work shows that these cues are important for closely-coupled interaction. Just telling a person what to do and how to do it is not very efficient e.g. for passing objects, demonstrating something or advising people. Cues like showing a direction are necessary for effective interaction. However, Hutchby also suggests that researchers should focus more on communicative affordances of the technology than cuelessness and psychological distance. Both may be considered worthy of investigation. Technology may be improved knowing the affordances condition on cues and distance.

Some argue that non-verbal signals not only constitute a separate channel of communication, but that they often override verbal content [Forgas, 1985]. In other words ‘how’ something is said can be more important than ‘what’ is said. One heavily used non-verbal cue is eye gaze, which is a richly informative behaviour in face-to-face interaction and serves at least five distinct communicative functions in conversation [Argyle & Cook, 1976; Kendon, 1990]:

- regulating conversation flow

- providing feedback
- communicating emotional information
- communicating the nature of interpersonal relationships
- avoiding distraction by restricting visual input

Comparable to turn taking in verbal communication, gaze is also an important precursor to interaction, regulating the beginning and ending of social encounters through the "making and breaking of mutual gaze" [Argyle, 1988]. Typically a speaker will make longer eye contact towards the end of his turn, often selecting the next speaker by ensuring that a mutual gaze is established with that person [Kendon, 1990].

Although facial expressions are very important, there are other non-verbal cues which should not be ignored such as gesture, posture and proxemics (personal space, discussed later in this chapter). Ekman divides gestures into different categories according to their communicative functions [Ekman et al., 1969]:

- *emblems* are used consciously and intentionally, and usually have a culturally codified meaning that can be substituted with a word or phrase, such as 'thumbs up'
- *illustrators* such as baton signals are directly tied to speech on a moment-to-moment basis, and are used to emphasise the rhythm of spoken dialogue
- *regulators* are used to mark the flow of the conversation as a whole, for example in indicating the next speaker with a hand gesture

One strong focus of previous research is the need to support *visual information*. Observations of face-to-face communication underscore that it is a complex multimodal process involving verbal and non-verbal communication [Goodwin, 1981; Clark & Brennan, 1991; Whittaker, 1995; Kraut et al., 2003]. This is especially true in complex object-centric tasks where visual behaviours play a central role for human interaction. In group communication settings involving objects, people jointly orient to, gesture at, and manipulate objects leading to observable changes of those objects. Non-verbal cues such as gaze, gesture and facial expressions are all highly reliant on visual information. This means being able to see each others behaviour is critical for the interpretation of gaze and gesture. Observing non-verbal behaviour enables participants to determine what objects other conversational participants are attending to, and what they are likely to talk about [Cooper, 1974; Huettig & Altmann, 2005].

In summary, non-verbal behaviours play a central function in face-to-face interaction. As Virtual Reality is a highly visual medium and CVEs allow unprecedented sharing of visual communicational resources, an avatars' [Schroeder, 2002] ability to convey such non-verbal cues is likely to affect how they are perceived as well as their contribution to social interaction.

2.2.3 Role of objects in communication

Interactivity of a medium is viewed as one of the key factors in facilitating the feeling of presence [Sheridan, 2000]. This could be explained with the use of an ecological perspective [Gibson, 1979; Flach, 1998]. The basic approach for this theory is that the environment offers situated affordances, perception-action coupling and tools become "ready-to-hand". The concept of affordances is associated with the work of Gibson in the psychology of perception [Gibson, 1979]. For Gibson, humans along with animals (insects, fishes, birds, etc.) orient to objects in their world (rocks, trees, rivers, etc.) in terms of what he called their affordances: the possibilities that they offer for action.

When a medium provides visual information about what a person is doing, the ability of people to ground utterances via actions sharply reduces the likelihood that they will provide verbal indicators of comprehension. Instead, they let their actions speak for themselves and demonstrate their understanding of the Helpers' utterances [Clark et al., 1991].

Verbal communication and body language are usually person-related, but for co-presence and interaction it is also important to consider the effects of none person-related communication. For example, the look and appearance of the environment is normally not person-related, however, it is more linked to the environment's desired task. Objects can be both person-related and non-related, and it depends in which context participants see them. For example, a business card is related to a person and a stone normally not. In addition, non-verbal communication through/with objects and the environment is of importance to enhance co-presence interaction.

Collaborative environments generally allow all participants to directly modify objects and to observe the effects of the changes made by others. Early studies of collaborative environments showed that adding this type of (non-verbal) visual information improves the efficiency of speech communication [Bly, 1988; Whittaker et al., 1993]. For example, Whittaker et al. compared speech-only communication with speech plus a collaborative environment for three different tasks: brainstorming, spatial design and collaborative editing. They showed that the environment improved communication for the latter two tasks, but not for brainstorming. Analyses of linguistic behaviour showed the reasons why: when the task requires reference to a complex layout (design and editing) or complex visual objects (spatial design), people were able to express complex spatial relations ("put that over here") and to use directed gesture. Participants were also more implicit in their communications when using the environment, because the environment supported situational awareness [Endsley, 1995]. Therefore, participants did not explicitly need to communicate changes about the current task if the collaborators could see this information directly. These effects were not found in the brainstorming task, which did not demand reference to complex objects, spatial relations or object transformations, something very common during closely-coupled collaboration.

These results about the primacy of objects are also supported by research into non-verbal communication. In conditions with few visual distractions the direct eye gaze rarely raised above 50% and looking at other people is the exception rather than the rule in conversation about things [Anderson et al., 1997]. In addition, gaze at others falls to 3-7% of conversational time when there are interesting objects present [Argyle & Graham, 1976] and mutual gaze is even lower [Anderson et al., 1997]. This suggests that participants do not spend entire conversations monitoring each other's facial expressions, especially when the environment contains relevant objects.

In the Virtual Gazebo application (see Chapter 4-7), objects are the main focus for interaction and one can observe that, for example, the concurrent sharing of an object is difficult due to technological and communicational shortcomings. Users are tempted to carry the object unsynchronised, which can be confusing. This can mean that users lose sense of direction and therefore the use of verbal communication increases to compensate for it.

2.2.4 Role of the environment in communication

Knapp uses the argument that each environment is made up of the three major components [Knapp, 1978]:

- The natural environment such as geography, location, atmospheric conditions
- The presence or absence of other people
- Architectural and design features, including movable objects

Some believe that "we are a product of our environment and that if we want to change our behaviour we need to only control the environment in which people interact" [Skinner, 1971]. The success of collaboration depends among other things on how people communicate or interact (including through objects) with each other. The environment and people's perception of it also play a vital rule. Knapp defined different perceptions of our surrounding [Knapp, 1978]:

- *Formality*: our reaction to the surrounding environment
- *Warmth*: nice colours, panelling, carpeting and furniture helping us to relax and to feel comfortable
- *Privacy*: an enclosed environment can give us the feeling of privacy and therefore we are more likely to interact closer and personal
- *Familiarity*: in a new environment we normally look for things that are familiar or speak with others about it
- *Constraints*: our reaction to an environment is also depending on whether we can leave it or not
- *Distance*: our response to a given environment can also be influenced by the Similar rules that are used in architecture for the design of buildings, homes or places

When designing an environment these perceptions have to be considered, otherwise people will feel uncomfortable and in a working environment this could lead to a performance reduction [Davis, 1984]. The same applies when designing virtual environments for an interactive task. The Virtual Gazebo application (see Chapter 4), as an example, used a deposit for the material and tools. And textures were used for wood and grass, which increased the feeling of lifelike, realism and made work more intuitive. For instance, with a single colour for grass, users had problems to recognise objects or to pick them up (the contours of the object appeared blurred). A real looking grass texture solved this problem. Similar things need to be done for the whole environment if needed. The application contained also more realistic looking trees around the building site to increase perception of being in a realistic environment. All these additions are not necessary for the task but they help to work there and to perform the task more efficiently.

2.2.5 Proxemics or Personal Space

The physical space that surrounds every person is used to define interpersonal interaction. In sociology the study of the construction of this space is called proxemics. People cannot make it through the day without referring to how one should line up or perhaps what a room should look like. Throughout the literature on proxemics, the most prominent author is Edward T. Hall as he coined the term "proxemics" [Hall, 1968].

An important aspects of Hall's studies of proxemics, is his notion and belief that proxemics are developed culturally, meaning that one's perception of personal space comes from one's culture and is therefore commonly shared with that culture. Hall has defined proxemics in several ways, but has always concluded that it is "how man unconsciously structures microspace – the distance between men in the conduct of daily transactions…" [Hall, 1963]. Hall's definition of proxemics goes on to include the organization of buildings and towns. Proxemics are usually distinguished into:

- *physical territory*, such as why desks face the front of a classroom rather than towards a centre isle, and
- *personal territory* that we carry with us, the "bubble" of space that you keep between yourself and the person ahead of you in a line.

Hall pointed out that social distance between people is reliably correlated with physical distance, and described four distances for personal territory (Figure 2-1):

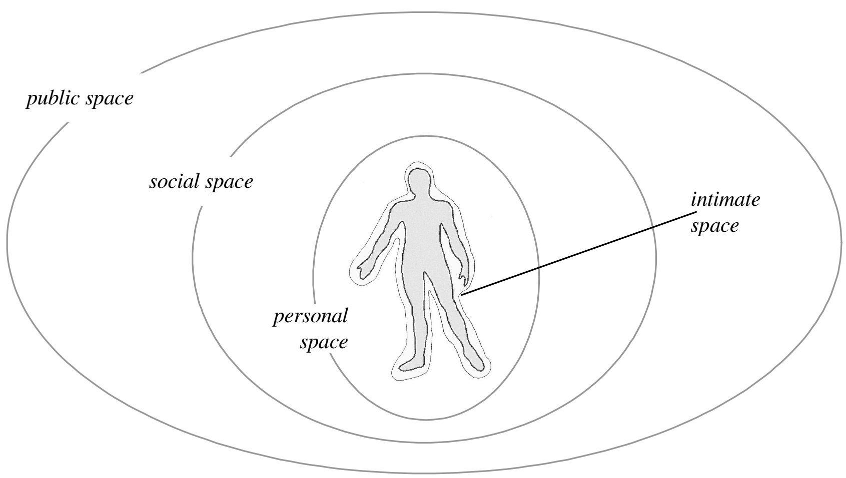


Figure 2-1: *Hall's model of four distances for personal territory*

- *Public space* ranges from 3.5 to 7.5 metres and is the distance maintained between the audience and a speaker.
- *Social space* ranges from 1.2 to 3 metres and is used for communication among business associates, as well as to separate strangers using public areas such as beaches and bus stops. At this distance, speech and facial expressions are clearly perceived, so communication can be efficient and accurate.

- *Personal space* ranges from 0.6 to 1.2 metres and this bubble of personal space generally surrounds people in their interactions with those known to them: it is a comfortable space within which people can discuss personal matters. People from different cultures, however, have a different perception of this personal space and usually this personal-space bubble is not evenly spread out all around but is rather larger in front of you.
- Finally, *intimate space* ranges some 30 centimetres and involves a high probability of touching. People reserve it for whispering, embracing, lovers, family, small children and very close friends.

It can be seen from the above that the sense of personal space and privacy is an integral element in human behaviour and interaction. When environments are designed without keeping these essentials in mind, those using the environment - the residents, staff, clients or inmates - are forced to operate in ways that make them uncomfortable. They may not understand or even perceive the reasons for their unease, but research shows that their discomfort will manifest itself in strained interactions and relationships [Painter, 1991].

In the design of virtual environments it is important to acknowledge this notion of personal space. People's behaviour and perception of proxemics do not necessarily change if they enter a virtual environment, especially when entering an immersive environment. The more people are immersed and have a feeling of presence the more likely it is that they behave as in the real world [Bailenson et al., 2001; Sander, 2005]. In addition, proxemics will influence people's perception of presence [Bailenson et al., 2004]. The use of CVEs is usually connected with collaboration with others and therefore people's proxemics will be violated in some way. For example, close interaction using immersive displays (see Chapter 3) usually occurs (using Halls definition) in the area of one's social and personal space. In contrast to the real world the intimate space can be violated in form of someone moving through the personal avatar. In a study by Slater et al. [2000a], subjects were found to be truly upset when their avatars were seen to pass through each other (due to the absence of adequate collision detection). Results by Bailenson indicated that participants maintained greater distance from virtual humans when approaching their fronts compared to their backs. In addition, participants gave more personal space to virtual agents who engaged them in mutual gaze and moreover, when virtual humans invaded their personal space, participants moved farthest from virtual human agents [Bailenson et al., 2003]. Similar behaviour is discussed in Chapter 6.7 (pp. 113) where users make an effort to move around the avatar of their collaborator or avoid moving through objects.

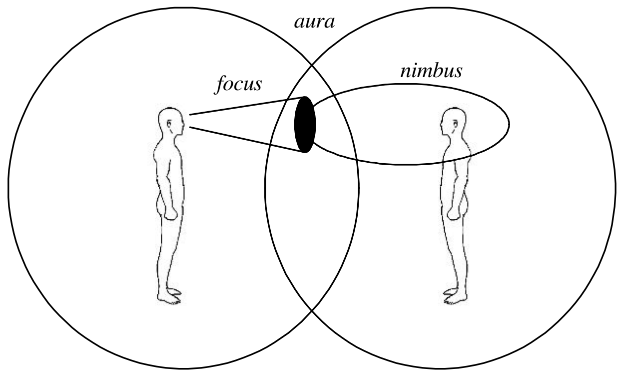


Figure 2-2: *Aura, focus and nimbus around avatars*

The management of awareness in virtual environments is related to proxemics in perception of space. The surroundings in CVEs can be cluttered at times and to improve performance awareness management can be used to render only close objects or transmit only audio for close people (spatial 3D audio). To manage awareness, avatars or objects are usually surrounded by an ever following and invisible aura [Benford & Fahlén, 1993]. It is defined as a sub-space which effectively bounds the presence of an object within a given medium and which acts as an enabler of potential interaction. Awareness is represented by a focus and nimbus allowing a user to interact with its surrounding (Figure 2-2):

- The more an object is within your focus, the more aware you are of it.
- The more an object is within your nimbus, the more aware it is of you.

2.2.6 Cultural differences of Communication

Communication may be a universal concept but large differences exist in the usage of communication between different cultures and the two genders. Intercultural communication has been of interest to communication researchers since the 1960s and is occurring "whenever a message producer is a member of one culture and a message receiver is a member of another" [Samover & Porter, 1985]. Most literature often includes discussion of subcultures, "a racial, ethnic, regional, economic, or social community exhibiting characteristic patterns of behaviour sufficient to distinguish it from others within an embracing culture or society", or co-cultures as an alternative term for subcultures so as not to imply inferiority in relation to the dominant society [Samover et al., 1985]. Other define co-cultures as "groups of persons united by a common element who live in a culture operating within a dominant culture" [Pearson & Nelson, 1994]. Scholars in the area of intercultural communication often begin their discussions by identifying the main characteristics of it. For example, Penington calls such elements "significant cultural components" [Penington, 1985], whereas Samovar and Porter identify what they call the "constituent parts of intercultural communication".

Three elements of intercultural communication common to most discussions are now briefly described for the purpose of illustration. Worldview, language, and non-verbal communication are often identified as important elements of intercultural communication. Worldview refers to a "culture's orientation toward such things as God, humanity, nature, the universe, and other philosophical issues that are concerned with the concept of being" [Samover et al., 1985]. An example often used is a comparison between Euro-American and Native American relationships to nature. While the Euro-American views the world as human-centred, the Native American views the human relationship to nature as being at one with nature.

Language is another significant element of intercultural communication, through which a culture expresses its worldview. Like culture in general, language is learned and it serves to convey thoughts, it transmits values, beliefs, perceptions, norms and so on [Penington, 1985]. The importance of language to intercultural communication is most obvious when cultures speak different languages. However, when each culture uses the same language, differences in meaning from culture to culture can be just as significant. For example, if a native American ask a Britain for the way to the restroom, the Britain may not know that he was asked for the way to the toilet. While this is an obvious example, Porter and Samovar point out that objects, events, experiences, and feelings have a particular label or name solely because a community of people have randomly decided to name them so. Language serves both as a mechanism for communication and as a guide to social reality [Samover et al., 1985].

Non-verbal communicative behaviour is another element that differs widely from culture to culture. For example, proxemics recognizes that "people of different cultures do have different ways in which they relate to one another spatially" [Samover et al., 1985]. Furthermore, the use of space helps define social relationships and social hierarchies. For example, it is known that a supervisor will exhibit a more relaxed posture than a subordinate, or that Arabs stand very close when talking while northern Europeans stand further away.

Some points of difference suggest that gender communication is a form of intercultural communication such as differences in worldview, language usage, and proxemics between the genders. In fact, it is difficult to discuss differences in worldviews regarding gender without talking about language, since our view of the world is expressed through language and other symbol systems. Tannen in her book argues that "communication between men and women can be like cross cultural communication, prey to a clash of conversational styles" [Tannen, 2001]. In part this is due to differences in the way men and women generally look at the world. For example, it is no coincidence that men use talk to exert control, preserve independence, and enhance status while women see talk as the essence of a relationship [Wood, 2004].

These are but a few examples of the ways in which differences in communication between the genders and cultures fit categories of primary elements in intercultural communication. The point is that these differences can create problems in communication and interaction. In intercultural communication, identifying problem areas can also help us learn to avoid them. These problem areas can be applied to gender communication as well. Barna identifies six stumbling blocks in intercultural communication: [1] assumed similarity, [2] language, [3] non-verbal misinterpretations, [4] preconceptions and stereotypes, [5] tendency to evaluate, and [6] high anxiety [Barna, 1985]. This last stumbling block, high anxiety, occurs when people are completely separated from their own culture. Awareness of the other five stumbling blocks, however, can be useful in improving our intercultural as well as gender communication.

Most CVEs don’t yet acknowledge those differences as most are designed as a proof of concept for virtual collaboration and to satisfy the research purpose. It should be noted, that all applications demonstrated and evaluated in this document have not acknowledged this issue either. They were designed for the purpose of studying display influences on collaboration and evaluated with people of similar ethnical background (European).

2.2.7 Trust in Interaction and Communication

When getting acquainted in face-to-face encounters, people tend to consider certain aspects of a person's appearance such as looks, body-language and other external signals which can tell them more about whom they are getting familiar with, and if the person in question is trustworthy or not. This first step in an interpersonal relationship towards a certain and trustworthy contact is taken in pure hope. One doesn't know the other person and has no experiences of him or her to fall back on. One lacks the so-called hard evidence that this person is trustworthy and is simply thrown upon one's optimistic hope that the person's appearance will correspond to one's expectations [Holmes & Rempel, 1989]. Sometimes the fragile hope about a deepened relation gets strengthened after this first examination and sometimes it expires entirely.

In general, trust refers to an aspect of a relationship between two parties, by which a given situation is mutually understood, and commitments are made toward actions in favour of a desired outcome. This means that trust is the willing acceptance of one person's power to affect another. No matter what type of long term relationship, people can build bridges by working together which in turn creates trust.

When virtual teams collaborate they need to establish trust at a distance in order to work efficiently. A study comparing four different communication situations (face-to-face, video, audio, and text chat) found that the three multimodal conditions were significant improvements over text chat [Bos et al., 2002]. It demonstrated also that video and audio conferencing were nearly as good as face-to-face, but showed some evidence for *delayed trust* (slower progress toward full cooperation) and *fragile trust* (vulnerability to opportunistic behaviour).

Some results indicate that subjects are more likely to deceive, be less persuaded by, and initially cooperate less, with someone they believe is in a distant city, as opposed to in the same city as them [Bradner & Mark, 2002]. Although people initially cooperate less with someone they believe is far away, their willingness to cooperate increases quickly with interaction. Bos et al. found that the co-located people formed an *in-group*, excluding the isolates which were remotely connected [Bos et al., 2004]. But, surprisingly, the isolates also formed an in-group, mainly because the co-located people ignored them and they responded to each other. This confirms that developers need to be concerned with developing technologies and avatars for bridging social distance, as well as geographic distance.

People also tend to pay attention to other avatar's movements and positioning in the virtual space. As Becker and Mark noticed, people in ACTIVE World (Action Control Training In Virtual Environments) are provoked if another avatar comes too close [Becker & Mark, 1998] or that people frequently make a comment if an avatar moves around with great speed and remarkably sweeping gestures. Becker et al. also notice that with the desktop interface, body gestures are seldom used in ACTIVE World, while emotions in general are communicated with text. So, when forming an opinion of a new acquaintance and estimating how trustworthy or reliable a person is, the physical appearance is, one could argue, of lesser importance when most ACTIVE World users presumably know that it has little to do with the person's identity and trustworthiness. On the other hand, social categories play a more important role when estimating how trustworthy other people are in virtual environments like ACTIVE World.

2.2.8 Communication in CVEs

The sections above have discussed how people communicate in the natural world and how they psychologically react, but there are many problems that need to be looked at before people can interact in a VE as they would in the real world. A VE allows people to overcome problems of remoteness and brings users together. Verbal communication in CVEs is mediated by audio connections between participants, whereas non-verbal cues are mediated through an avatar, the virtual representation of a participant. The quality and degree of these cues can vary depending on the communication support by the application and technology. Chapter 3 will investigate and discuss various tele-operation systems with focus on communication support.

Collaboration between people sharing the same workspace involves the ongoing and seamless transition between individual and collaborative tasks, shifting their attention from a state of peripheral awareness to focussed awareness. In CVEs peripheral awareness is limited to the field of view and the happenings in that field of view and further it is possible to perceive others without being perceived, for example using the bird's-eye view. This can lead to a break down of the understanding of each others perspective. Earlier work on human interaction looked at the way in which people use their viewpoints and react to gestures [Hindmarsh et al., 2000]. Hindmarsh et al. observed that a desktop user, when directed to an object by gesture and verbal comment, tends to visually locate the user and then follow his gesture to locate the object. This can lead to confusion when the directing user is changing position or gesture. Attempts were made to reduce these problems by using “out of body camera view” [Greenhalgh & Benford, 1995], two additional windows into the virtual environment or peripheral lenses. Although the latter solution enhanced the awareness, results show that peripheral lens distortion can disrupt both a user’s own sense, and their notion of the other’s sense, of orientation to actions and features within the environment [Fraser et al., 1999]. In contrast, fully immersive technology (see Chapter 3) places a user in a spatial social context allowing natural first person observations of remote users interacting with objects. This improves the work and awareness within such an environment (see Chapter 6), and when connected with other non-immersed users, it can be observed that, in addition, the immersed user adopts a leadership role ([Steed et al., 1999; Slater et al., 2000a], Chapter 5).

Communication about a task in the real world will refer to the artefacts used as part of that task, using a variety of movements, such as changes in bodily orientation and gestures. A rich CVE interface, with effective virtual embodiments and easy to use navigation controls should help participants to reach the working consensus for the interaction, especially in those situations where the role structure is initially not obvious to the participants. According to some research the most fruitful collaborative behaviours for understanding and observing social interaction are head nodding, face looking, smiling, head touching, and speaking, including simultaneous speech [Argyle, 1988]. In immersive CVEs these features a routinely implemented as well as used and Chapter 6 will discuss the fluent workflow this can create.

2.3 Summary

The term presence is not clearly defined but commonly referred to as a feeling of “being there other than the current physical environment” alone, with others, objects or in a social setting. Research shows that presence influences task performance but is itself influenced by a variety of factors [Singer et al., 1995; Snow, 1996; Welch, 1999; Nichols et al., 2000; Slater et al., 2000b; Schroeder et al., 2001; Zimmons et al., 2003]. The experience of presence is a precondition for co-presence [Slater et al., 2000a] which in turn is a precondition for collaboration [Tromp et al., 1998]. Improving presence therefore improves collaboration.

Although presence and co-presence is difficult to measure, some subjective measurement is possible, typically using questionnaires, whereas objective measures include psychological measurements and observed reactions to given stimuli. For example, when an immersed user ducks down in order to avoid a collision with a virtual beam suspended in midair (e.g. see Chapter 6). Results suggest that with the intensity of collaboration, a higher perception of co-presence can be found. Chapters 5-8 analyse various displays and results show that presence is depending on display properties such as field of view, size and navigational freedom.

The feeling of presence, and particularly the naturalness of interaction with objects, may be improved when the user can see their own body in the context of the virtual environment [Mine et al., 1997]. Schuemie concludes that little is known about what interaction has to do with presence [Schuemie et al., 2001]. It may be argued that even less is known about the relationship between effective interaction on common objects as a focus of interest and co-presence. An understanding of the nature of interaction in the real world can help to reason about co-presence and may lead to further defining its requirements.

While we collaborate with other people through an object, we use a variety of communicational resources to demonstrate our opinion, intention and needs to others. Be it simply verbally with emotional nuances, with gestures and postures in a non-verbal way or by manipulating the object directly. When interacting remotely, these forms of social human communication (SHC), as well as the representation of the object, need to be mediated through tele-collaboration technology. For example, in CVEs this is done by using avatars which are a virtual representation of a user and immersive technology place a user in a spatial social context allowing natural first person interaction. The avatars used in most systems are far from sophisticated or lifelike, yet people accept them and when working closely around shared objects and other artefacts the avatar appearance is becoming less important [Nilsson et al., 2002]. Lifelike avatars would be very useful for a social discussion if all minor cues are represented, but for an object oriented task it is more important to communicate the intentions across.

Conversation and collaboration is subject to alterations also called turn-taking and its problems have been widely discussed in research on remote communication and collaboration, but the key problem is that it is very difficult to express ideas, emotions and opinions. More effort is needed than during face-to-face meetings and people have to be very explicit in what they say and how they refer to objects and their actions. If supported, the use of postures, gestures and other non-verbal cues can simplify the communication between groups.

In the design of virtual environments it is important to acknowledge the notion of personal space. People's behaviour and perception of proxemics do not necessarily change if they enter a virtual environment, especially when entering an immersive environment. It can be observed that an immersed user in collaboration with other users does avoid violating their personal space. For example, one user is holding an object while a second user comes with a tool to manipulate the object, but avoids entering the intimate space.

The subject of communication is not always abstract and often relates to our surroundings and artefacts within it, both providing a context for understanding and people are fully adaptable to this. Differences can arise when we meet people of other cultures or use technology for communication and collaboration.

Chapter 3

3 Supporting Remote Collaboration while keeping workflow

Today talking to other people from remote places is ubiquitous using mobile phones and researchers are just beginning to exploit how natural interaction with artefacts can be made possible and just as ubiquitous from distributed sites. People often use objects when working alone or within a group. It is therefore natural that researchers and developers try to achieve such interactivity with distributed technology. Yet a number of technological issues have to be overcome in order to create a system which allows for easy, straight forward distributed shared object interaction, as simple as using a mobile phone.

Chapter 2 introduced the notion that (Co-)Presence and SHC are very important for good collaboration between people. Subconsciously as well as consciously people recognise nuances from gestures, postures and verbal cues. Furthermore, the feeling of presence can enhance the performance of a team and help to recognise SHC cues. Chapter 2 also introduced briefly the notion of a fragmented workflow on the example of desktop CVEs. Hindmarsh et al. [2000] observed that a desktop user, when directed to an object by gesture and verbal comment, tends to visually locate the user and then follow his gesture to locate the discussed object. This focus on common ground can take considerable time (>20sec) and thereby fragment the workflow. CVEs are only one example of technology that allows people to communicate and collaborate remotely. This chapter will discuss how other technologies allow distributed collaboration and why most of these fail in removing or at least reducing the fragmentation problem. It will conclude with a summary of different participation frames that each technology supports. The latter is a distinction on how distributed object collaboration is supporting either look-into, reach-into or step-into someone's environment. But first parameters for the following discussion shall be defined.

3.1 Criteria for Evaluation

Collaboration between a geographically separated group relies on technology and the challenge is to understand and reduce restrictions and limitations introduced by this technology. Different technologies are distinct in the way they convey interaction, attention and awareness to others, shared objects and the environment. Thus, the main criteria for evaluating these technologies will be their support for naturalness of communication between people and object manipulation without fragmenting the workflow. Furthermore, chapter 2 discussed that co-presence has a positive effect on collaboration, but the following evaluated technologies have different support for creating this togetherness.

3.1.1 Naturalness

Interaction with an environment, objects or others is determined by effective means of manipulation and communication. In some cases this could be enhanced through abstract interaction techniques or pre-defined work processes (e.g. using the mouse or macros while writing this document). However, in a collaborative setting it is more likely that performance is increased by natural and intuitive interaction, even though mediating and supporting this through tele-collaboration technology is challenging. This is because natural interfaces allow people to concentrate on the task instead of compensating for problems with the interface. Lok et al. investigated how handling real objects and self-avatar visual fidelity affects performance on a spatial cognitive task in an immersive VE. Their findings show that interacting with real objects significantly improves task performance over interacting with virtual objects [Lok et al., 2003]. While designing interfaces for people collaborating via technology, close attention must be paid to how users think and perceive. The information must be clearly visible; interaction techniques should be simple (natural) and account for both novice and expert users.

3.1.2 Being together (co-presence)

During collaboration with others a sense of togetherness can be felt, which is usually referred to as a feeling of co-presence. This perception of spatial and social togetherness between remote people when collaborating is further enhanced while collaborating around shared objects [Durlach & Slater, 2000]. Various factors contribute to this sensation [Slater et al., 2000a; Schroeder et al., 2001; Mortensen et al., 2002] and chapter 2 has discussed these in more detail. Technology is certainly such a factor and therefore its effects should be considered in this evaluation of tele-collaboration technologies. For example, the way technology supports spatial proximity (Chapter 2, [Benford et al., 1998]) can influence our awareness of others as well as objects. In addition, the lack of social cues reduces awareness which affects collaboration, communication, presence and co-presence [Bowers et al., 1996]. In a collaborative scenario awareness is created by sharing both a social and spatial context between collaborators as well as sharing objects and the environment.

3.1.3 Fragmentation of Workflow

Disruption in the work process for re-orientation within the scene can have implications when working together. This could go as far as to the point that collaboration is interrupted, because teamwork relies on a fluent and coherent multimodal communication between the partners that can be disrupted by the technology. For example, a person may reference an object by speaking its name and pointing to it, but the technology may fragment the workspace (e.g. due to limited FOV) such that the meaning of the gesture is lost or fragment the nuance such that the gesture occurs at a different time to the spoken description. It usually means that the partner needs to re-orientate and interrupt his work in order to see the other partner and the object they are working with [Hindmarsh et al., 2000]. This fragmentation of the workflow can be time-consuming and therefore have a negative impact on performance. A subsequent study tried to resolve some of the issues with peripheral lenses, which resulted in an enhanced field of view (FOV). Although this solution enhanced the awareness, it also showed that peripheral lens distortion can disrupt both a user's own sense, and their notion of the other's sense, of orientation to actions and features within the environment [Fraser et al., 1999]. Chapter 6 will demonstrate that immersive display can reduce this fragmentation and lead to a seamless communication and interaction between collaborators.

To maintain a continuing workflow disturbance from the surroundings should be avoided or decreased. For example, a break in presence [Brogni et al., 2003a; Brogni et al., 2003b] can interrupt the workflow and reduce the feeling of being together. Some systems embody remote users through computer graphic virtual characters, known as avatars. Avatars reflect the orientation of the user within the environment and in some cases gestures that can be directed towards the focus of interest. They therefore provide a basic means for object focussed social interaction [Heldal et al., 2005]. The impact of avatar representation and gestures on collaboration has been well studied [Benford et al., 1995; Bente & Kraemer, 2002]. Ensuring that remote participants can easily see what each other are really looking at and doing is important and this thesis will later show (see Chapter 6) that this can significantly impact on the flow of conversation and work.

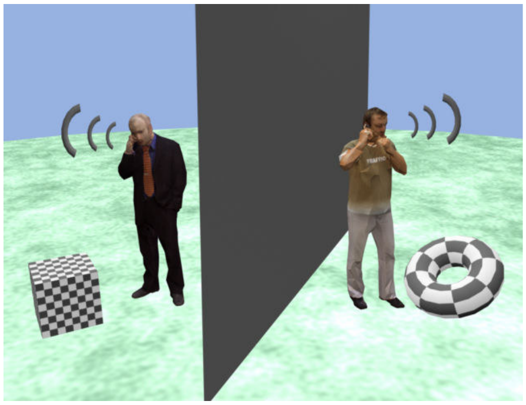
The above demonstrated that continuous working can be interrupted due to technological shortcomings or external influences. While the latter includes social talks with colleagues, other required work tasks and so forth, these influences are difficult to control and occur naturally. In contrast, technological interruptions are not necessary, which is why technology should be continually improved. Unfortunately, technology does not lend itself for every task and the following sections will discuss various technologies in their support for tasks involving remote collaboration with shared objects.

3.2 Audio Conferencing

The development of the telephone allowed people for the first time in history to communicate directly with each other from remote places with instant reply. However, chapter 2 discussed how human communication is more than just transmission of speech, because of a complex multimodal[1] process involving verbal and non-verbal communication [Clark et al., 1991; Whittaker, 1995; Kraut et al., 2003]. As audio conferencing is a unimodal technology, it is known to show limitations in supporting essential social cues [Short et al., 1976; Mark et al., 1999]. Gestures and other non-verbal cues are not transmitted and participants have to establish a mechanism for turn-taking [Hutchby, 2001]. Cultural differences can add to the problem as correct mechanisms become essential to avoid offences. Nevertheless audio-conferencing is known to be very effective for brainstorming tasks [Whittaker et al., 1993; Olson et al., 1995]. The sharing of objects and data is complicated since a common reference is missing and collaborators have to share a mental model formed through verbal discussion, but this can be flawed when people think about different things (Figure 3-1). For example, the simple task of describing the shape and look of a mug can be challenging. Common references, such as shapes, geometry and colour definitions, can help collaborators to form a shared mental vision of the mug, yet people may still imagine different things. In such a situation an image would solve this issue in an instant. Likewise, a conference setting where people are not necessarily aware of each other, and don't have eye contact nor can observe each other's reaction, communication can be complicated and the feeling of being together is reported low [Boyer et al.,

[1] provides the user with multiple modes of interfacing

1998]. This fragmentation of space leads to fragmentation of workflow and consequently has an effect on performance [Wainfan & Davis, 2004, pp. 34-39].

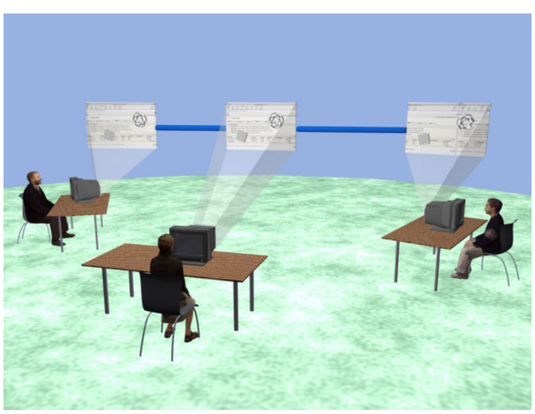


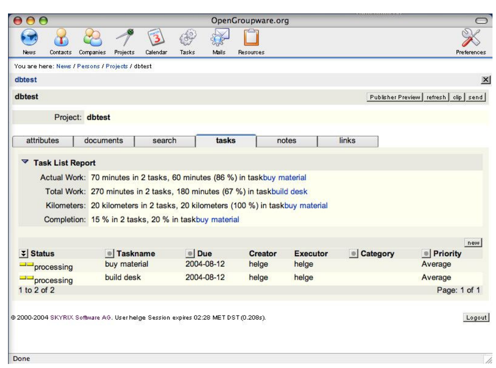
Telephone Conference

Figure 3-1: *Audio Conferencing (unimodal technology)*

3.3 Groupware

A very popular and common medium for remote collaboration forms a class of technology / application defined as groupware. Within this discussion the term groupware is being used to refer to windows-based collaborative applications such as CSCW solutions. In contrast to audio conferencing, communication is mostly an asynchronous manipulation via a desktop interface, with limitations in communication of verbal and non-verbal cues (Figure 3-2). It does not allow informal interactions or concurrent object manipulations that may be necessary to react to irregular actions with the team [Markus & Connolly, 1990; Antunes et al., 1995]. Yet it is widely used for cooperation between co-located as well as remote people. The market for groupware is vast and many applications exist and compete, of which OpenGroupware is a free open source example.





OpenGroupware.org

Figure 3-2: *Groupware*

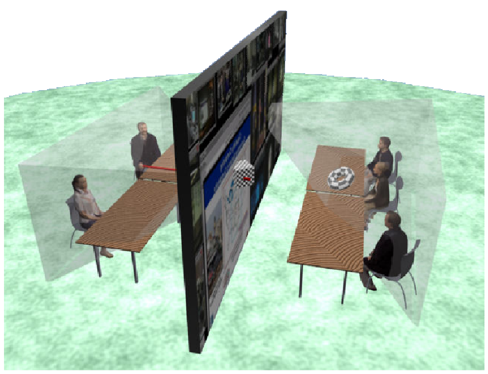
A limitation of groupware is in perceiving an awareness of action of others [Rodden, 1996; Berlage & Sohlenkamp, 1999; Schmidt, 2002]. For example shared interaction on the same document is usually prevented as a user has to lock out the document. This can be made known by the application to other collaborators and thereby trying to create an awareness of collaboration. However, as the document is locked other collaborators are unable to modify or add their own input reducing the level of togetherness. At the same time this is fragmenting the workflow as users have to work on a copy and changes need be merged with the latest edition after it is unlocked risking inconsistencies. A typical example is the writing of a conference paper, where multiple authors work on the same document. This can be a tedious and time consuming task as authors work on different document versions, introducing different ideas or work on the same section at the same time. In the early stages of a paper it would be useful to get everyone's input in order to create a document framework. Unless authors meet in a conference setting, this is usually an asynchronous development with many document iterations. The first author's responsibility is to collect the various inputs and to create a revised paper version, which in turn will be reviewed and extended by the co-authors. Groupware can help to keep track of the various versions and is used as a medium of communication. In a face-to-face meeting the progress on such a paper is enhanced as all authors have the chance to react immediately on suggestions and comment. This synchronicity is very difficult for groupware to achieve, which means that the work process is fragmented and collaboration usually decoupled.

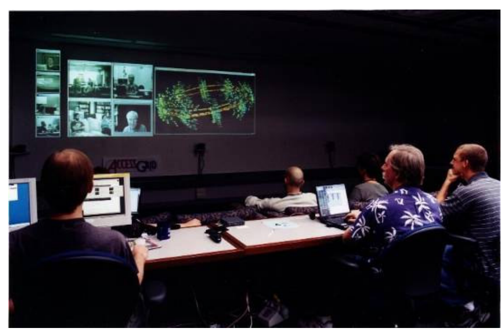
3.4 Video Conferencing

Video conferencing (VC) is the simultaneous exchange of video and audio data between remote sides. Compared to the previously discussed audio conferencing and groupware, VC introduces new limitations to the ordinary transfer of information. Just as with audio, jitter can interrupt a transmission resulting in communication problems. This, however, is more a problem for audio than video as humans are adapted to receive a constant audio stream while closing the eyelid results in a break of the "video" stream. However, video in contrast to audio requires a higher bandwidth for transmission which could lead to an increase in latency and may even result in jitter.

3.4.1 Tele-conference

Tele-conferencing is when two or more participants at different sites connect by using computer networks to transmit audio and video data. This means that non-verbal cues, such as gestures, can be transmitted via video. For example, for a point-to-point (two-person) tele-conferencing system each participant has a video camera, microphone, and speakers mounted on his or her computer. As the two participants speak to one another, their voices are carried over the network and delivered to the other's speakers, and whatever images appear in front of the video camera appear in a window on the other participant's monitor. Multipoint tele-conferencing allows three or more participants to sit in a conference room and communicate via a big video screen (Figure 3-3). Until the mid-nineties, the hardware costs made Tele-conferencing prohibitively expensive for most organizations, but that situation is changing rapidly and tele-conferencing is becoming more and more a common tool in global organisations.





Access Grid [Childers et al., 2000]

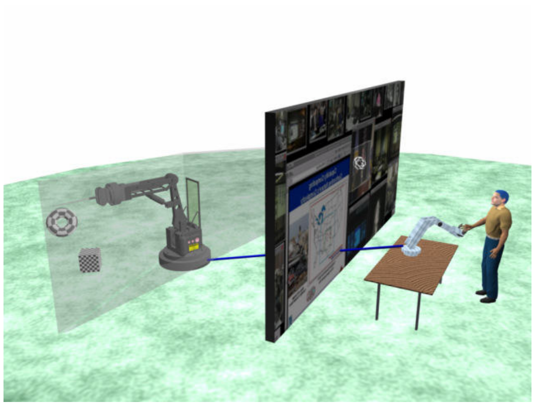
Figure 3-3: *Video Conferencing (Tele-conference)*

Tele-conferencing places remote people in a setting where they can look-into each others environment. The instant transmission of audio and video can be complemented with groupware elements such as shared white boards. This allows participants to manipulate virtual objects and data during a conference. A typical example of tele-conference application is AccessGrid [Childers et al., 2000]. Similar to groupware, concurrent manipulation is rarely supported and interaction techniques are limited to desktop computer interfaces. The sharing of physical objects is limited to interaction from one site and observation at the others. This can create delays in communication especially when misunderstandings occur. Moreover, research showed that adding visual information may impair critical aspects of spoken communication creating additional delays [Anderson et al., 2000]. For example, many videoconferencing systems introduce delays into speech by buffering it so that it can be synchronized with video. But several studies showed that such delays compromise important communication feedback processes that demand immediacy: e.g. backchannels, or interruptions [O'Conaill et al., 1993; Anderson et al., 2000; Whittaker, 2003]. This can affect the outcome of conversations. So by trying to visually enrich communication channels, communication itself can get disrupted. However, modern high speed connections can reduce such delays below 200ms which is not worse than a telephone call.

Large screens and wide-angle cameras can create a good projection of the other room thereby increasing awareness of and togetherness with others. However this is only true as long as participants can be seen. As soon as people move out of the camera frame their actions become invisible, because one only "looks into each other's world", which limits the operating range to move and to be seen. In addition, it is difficult for all participants to interact with a shared object. In particular, it is hard to see how someone is interacting with an object when the operator, observer and object are each in separate windows, as in Access Grid. Summarising, spatial separation between each site and with data as well as delays during communication leads to a fragmentation of workflow making tele-conferencing an ineffective tool for closely-coupled collaboration.

3.4.2 Tele-haptics or Tele-operation

Tele-operation is the principle means of controlling robot end-effectors and can be seen as an extension of tele-conferencing. The aim is to provide the user with a way to manipulate objects remotely rather than just a "look-into" environment [Buxton, 1992; Hollan & Stornetta, 1992]. The traditional way of tele-operating a robot has been using a data glove or exoskeleton master: there is a direct mapping from the human hand to the robot end-effector. Generally, the finger positions of the human master are translated to the robot and visual or force feedback are returned from the robot to the master [Burdea et al., 1992; Speeter, 1992]. Currently there are several difficulties with this approach regarding the calibration, autonomy and degree-of-freedom of the robot. Firstly, it is difficult to find a direct mapping from the human hand master to the robot due the complex nature of the musculoskeletal system involved in the human hand. Robotic devices are usually inferior in their capabilities compared to human hand range of motion and degree of freedom. Secondly, robot commands are displacements rather than functions. Without a high-level function, it is difficult to enhance a robot's autonomy. Furthermore, in situations in which there are long communication delays between the master and the robot (greater than 1 second), it is useful for the robot to perform certain functions autonomously [Bejczy & Kim, 1990]. And finally a high degree-of-freedom force feedback is still experimental and expensive.



Tele-operation [Ohba et al., 1999]

Figure 3-4: *Tele-operation*

Although tele-operation can help solving the problem of remote manipulation of real objects, it is currently not possible for multiple users to manipulate the robot or to share the object manipulation. In addition, the awareness and fragmentation problems experienced in tele-conferencing ("looks into someone's world") exist as well. The robot is merely an extension of the remote user's hand but the subtle nuances of non-verbal communication can't be transmitted yet and its main function is to interact with objects in the remote environment rather than as a direct communication medium. This makes tele-operation a first step to reach-into someone's environment (Figure 3-4).

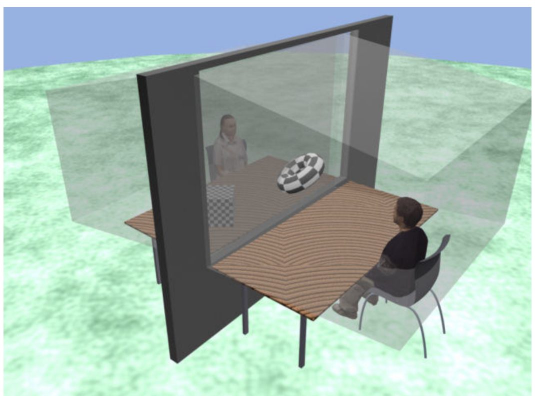
3.5 Mixed Reality

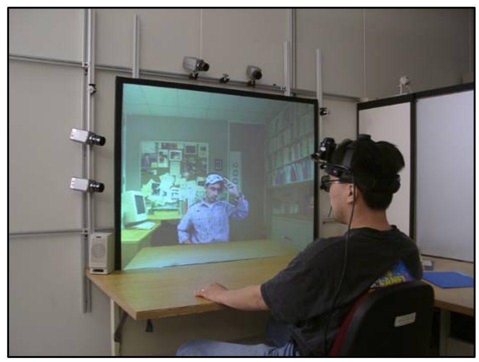
Mixed reality technology is a combination of traditional technology with computer generated virtual elements and it promises to enhance remote face-to-face communication. Various approaches are taken but all are trying to blend the physical and virtual worlds so that collaborators can interact with the 3D digital content and improve users' shared understanding.

3.5.1 Tele-immersion

Tele-immersion is defined as the integration of audio and video conferencing, via image-based modelling, with collaborative virtual reality in the context of data-mining and significant computation. Some argue that the ultimate goal of tele-immersion is not merely to reproduce a real face-to-face meeting in every detail, but to provide the "next generation" interface for collaborators world-wide, to work together in a virtual environment that seamlessly converges from the real local into the remote virtual environment [Leigh et al., 1999].

An advantage of tele-immersion compared to tele-conferencing is the seamless transition from a local to a remote environment. Recent implementations are reported to come close to a face-to-face meeting [Raskar et al., 1998]. Although they don't allow manipulation of real objects, it is possible to integrate virtual objects within the view field where these objects can be manipulated by the collaborators.





Office of the future [Raskar et al., 1998]

Figure 3-5: *Tele-immersion*

The technical complexity to create such a connection is, however, very high as multiple video streams have to be processed and transmitted to the remote location. The scalability is therefore rather limited as only a few places can be connected. Similar to tele-conferencing this technology allows people to look-into each others environment rather than to share it (Figure 3-5). This means that it is still difficult to change to any view angle or to point to an object in the other's workspace. The consequence of this setup is that shared object manipulation is restricted on to the shared desk space.

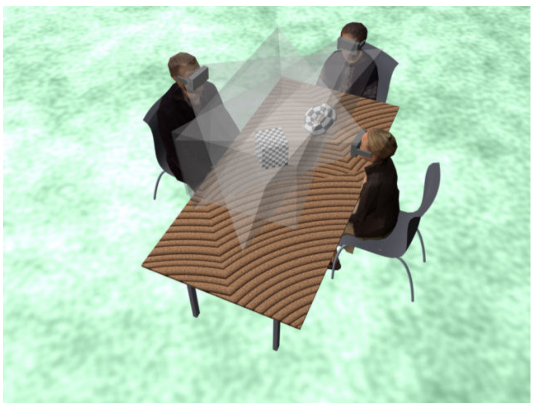
The awareness of each other's presence and the feeling of being together is, however, very high and probably the closest yet to a face-to-face meeting [Lanier, 2001]. Due to the high resolution projection and seamless transition from the real table to the remote virtual table, it is easy to pickup non-verbal cues and facial expressions [Kauff & O.Schreer, 2002]. However, as with most mixed reality displays, the user has to wear glasses in order to see additional content and this can handicap transmission of facial cues [Towles et al., 2002]. Nevertheless, compared to pure video-conferencing this technology has little effect on the workflow within its spatial constrains [Lanier, 2001]. For example, a face-to-face conference meeting would be easy to create with this technology with a high degree of co-presence. As this technology tries to extend from the real world into a remote environment with seamless transition its limits lie within the spatial constrain towards remote users. This means that it is not possible to get close to or even behind a user and its main application is an extension of conference tables. However, a number of collaborative tasks require free movements within the space and towards shared objects. In addition the scalability issues [Towles et al., 2002] of the technology hinder large setups and alternative technology is required for more flexible applications and simulations.

3.5.2 Collaborative Augmented Reality

Augmented reality (AR) is a technology which superimposes computer-generated images onto a user's perception of the real world via devices such as “see-through” or “see-around” displays (e.g. HMD). As a consequence, additional 3D information such as objects or other virtual characters become available to the user. Researchers have developed single-user AR interfaces enabling people to interact with the real world in ways never before possible. For example, surgeons can see virtual ultrasound images overlaid on a patient’s body or head-up displays may guide as part of a navigation system by overlaying route information without disrupting the view to the real world. This kind of navigation system might be very suitable for motorbike driver, integrated in a helmet, or car driver as part of the front screen.

One of the most interesting features of co-located AR is the seamless nature of collaboration. Users see each other and at the same time they see virtual objects between them (Figure 3-6). Unlike previously introduced technologies, co-located AR interfaces do not separate the communication space from the task space, allowing users to interact with virtual content by using familiar real objects and thereby avoiding interruptions in the workflow [Ulhaas & Schmalstieg, 2001]. In a study comparing AR conferencing to traditional audio- and videoconferencing, subjects reported a significantly stronger sense of presence for their remote counterparts in the AR conferencing condition, and that it was easier to perceive one another’s facial expression cues [Billinghurst & Kato, 2002].

A problem with current AR displays is that viewing the world through it is not the same as seeing it with the naked eye. Current see-though HMDs have limitations in their field of view, resolution, and colour depth. They allow users to view the world normally, but it is difficult to build see-through displays with a wide field of view. Another challenge is the problem of tracking and registration of objects and people. Users' viewpoints need to be tracked, in order for virtual models to be overlaid precisely on the real world. For example, computer-vision-based tracking techniques work only when physical tracking markers are in view and may introduce more system delays compared to other magnetic or ultrasonic tracking technologies.



Chess [Ulhaas et al., 2001]

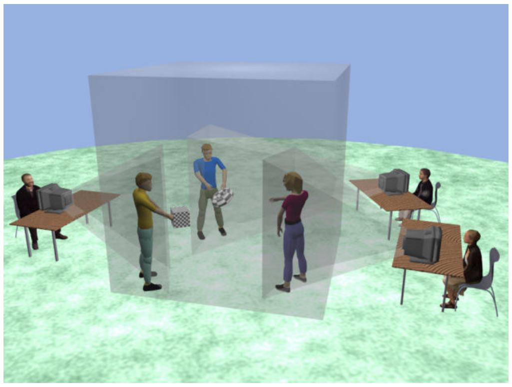
Figure 3-6: *Collaborative Augmented Reality*

Collaborative augmented reality enhances AR with distributed system support for multiple users allowing a co-located joint experience of virtual objects [Billinghurst et al., 2002; Schmalstieg et al., 2002]. However, technology for remote AR collaboration also involves further limitations. It is difficult for current technology to provide remote participants with the same experience they would have if they were in a co-located meeting. Some desktop based AR systems employ video conferencing technology with AR technology but the personal workspace is very constrained and scalability is low [Barakony et al., 2003]. Other more portable systems require headsets with low image resolution and limited field of view or instead employ handheld devices with small displays. In addition, the accuracy of the tracking and registration of people and objects in an unconstrained environment is problematic. Sometimes Global Positioning System (GPS) is used which provides flexibility but has a low resolution and no orientation. Hybrid tracking approaches combining several techniques seam to be a promising direction [Billinghurst et al., 2002]. However, with the tracking problems mentioned, the full support of non-verbal communication is limited and constrained. Furthermore, differences in real world environments between remote collaborators make it difficult to share space and therefore workspace references only function if all participants work in identical environments or references are made only to virtual objects.

3.5.3 Non-immersive Collaborative Virtual Environment

Collaborative Virtual Environments (CVEs) are computer-enabled, distributed virtual spaces or places in which people can meet and interact with others, with agents and with virtual objects. CVEs vary greatly in their representational richness from 3D virtual reality to 2D and even text based environments. The main applications to date have been multiplayer games, social meeting places, military and industrial team training, engineering, collaborative design and medical treatment [Pausch et al., 1992; Becker et al., 1998; Prasolova-Førland, 2005]. When CVEs were first developed in the 1990s they were seen as cheap alternatives to video conferencing and teleconferencing. In reality however they can be far more effective for remote collaboration: most significantly, they represent a shift in interacting with computers and communications technology in that they provide a space that contains or encompasses data representations and users [Snowdon et al., 2001]. Tele-conferencing does not provide full body language or other spatial cues such as gaze direction, spatial presence and direct or peripheral awareness of the activity of participants. In addition they are also weak in terms of shared activity awareness [Fussell et al., 2000] and are not suitable for highly scalable and distributed deployment.

It is important that collaborators know what is currently being done and what has been done in context of the task goals. Individuals need to negotiate shared understandings of task goals, sub-task allocation and of task progress. By making the actual work take place within a CVE, collaborators can be aware of each other's activities. Shared objects become not only the subject of communication, but also the medium of communication: as one user manipulates an object, changes are visible to other users [Benford et al., 1997; Snowdon et al., 2001].



Desktop Gazebo [Roberts et al., 2003]

Figure 3-7: *Collaborative Virtual Environment*

In tele-conferencing and in other mixed reality implementations, the visual channel usually shows a ‘head and shoulder’ view of other participants, providing information about their gaze and facial expressions. This “talking heads” view contrasts with the visual information presented in CVEs. Instead, CVEs provide visual information about relevant shared objects (such as documents or drawings) that the participants are jointly working on as well as the participants themselves in a spatial context to these objects (Figure 3-7). Thereby it also allows all participants to directly modify those objects and to observe the effects of changes made by others.

CVEs clearly have the potential to enable innovative and effective distance teaching techniques, involving for example debate, simulation, role play, discussion groups, brainstorming, and project-based group work. The emphasis can be placed on the human-to-human interactions as common understandings are negotiated and developed across differences of knowledge, skills and attitudes. The increased sense of social presence means that participant’s absence or non-participation is less likely to go unnoticed. In addition, in its nature the sharing of objects is easier to achieve than in other technologies. Even though a study by Broll in the mid-nineties concluded that concurrent shared manipulation of objects in a CVE would not be possible with technology at that time, due to delays caused by distribution [Broll, 1995]. Such interaction, however, was demonstrated in the following years, at first using single desktop system (e.g. [Ruddle et al., 2002]) and later through networked immersive displays (e.g. [Mortensen et al., 2002; Linebarger et al., 2003; Roberts et al., 2003]). Some results and evaluations are presented and discussed in Chapter 5-8.

Desktop based CVEs still provide a look into environment, however, they share objects within a shared space. This means that user movements are poorly mapped to the virtual world and the support for non-verbal communication is rather limited due to the lack of natural avatar representation [Ståhl, 1999; Robinson et al., 2001]. Participants may use keyboard commands to create gestures if they don't find this method too tedious to use. However, as users share the same environment through their avatar, they are able to separate themselves from groups, for example, for private conversations with another participant. People in these environments are also extremely aware of their personal space and react with hostility to someone violating it [Bailenson et al., 2003]. This has an influence on a users perception of presence [Bailenson et al., 2004], whereby the notion of having some sort of virtual representations of participants in a collaborative virtual environment is in general very important to create a sense of presence, especially co-presence [Slater et al., 1993; Benford et al., 1995; Durlach et al., 2000]. In addition, both notions are perceived to be higher if the task is highly collaborative [Casanueva et al., 2000].

Hindmarsh et al. studied collaborative interaction of two users through a set of objects using a desktop based CVE [Hindmarsh et al., 2000], in which the participants were asked to rearrange furniture. The authors found that the limited field of view (FOV) on desktop systems was of great hindrance due to problems with fragmentation of the workspace. It took an unnaturally long time (>20sec) for users to perceive each other's gestures and to reference them to the places and objects in their conversation. The authors concluded that this was caused from a lack of information about other's actions due to their limited window into the world. In addition the study found problems with slow applications and clumsy movements as well as the lack of parallelism for actions. A subsequent study tried to resolve some of the issues with peripheral lenses, which resulted in an enhanced FOV. Although this solution enhanced the awareness, it also showed that peripheral lens distortion can disrupt both a user's own sense, and their notion of the other's sense, of orientation to actions and features within the environment [Fraser et al., 1999].

3.5.4 Immersive Collaborative Virtual Environment

The term "immersion" describes the extent to which a given technology replaces real world stimuli with synthetic stimuli within the virtuality continuum [Milgram et al., 1994]. A necessary condition is Ellis' notion of a Virtual Environment (VE) as a communication media [Ellis, 1996], maintained in at least one sensory modality (typically the visual). The degree of immersion is increased by increasing the field of view (FOV), greater degree of body tracking, decreased lag between body movements and resulting changes in sensory data amongst others [Pausch et al., 1997; Sheridan, 2000; Baños et al., 2004]. Immersion may lead to a sense of presence (see Chapter 2) and is a necessary rather than a sufficient condition for presence - immersion describes a characteristic of the technology, whereas presence describes an associated state of consciousness [Slater et al., 2000b].

Traditionally CVEs were used on desktop interfaces, but more recently a growing critical mass of studies have used Immersive Projection Technology (IPT). There exist two classes of IPTs: fully immersive and spatially immersive. The difference is that in fully immersive displays (like head-mounted displays (HMD)) the user is completely cut off from references to the real world, while spatially immersive displays like 4-sided CAVE [Cruz-Neira et al., 1992] still allow for outside references. A number of studies have highlighted pros and cons comparing HMDs to CAVEs and emphasised the behaviour of participants using both kinds of display devices [Pausch et al., 1997; Coomans & Timmermans, 1998; Schubert et al., 1999; Kjeldskov, 2001; Bowman et al., 2002a; Manek, 2004]. Figure 3-8 to Figure 3-10 show the various immersive displays used for the research in this thesis.

Trimension Reactor
- resolution of 1024x768
- FOV 160º- 270º
- magnetic or ultrasonic tracking

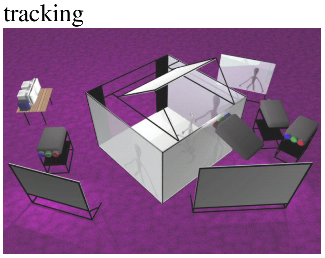


Figure 3-8: *CAVE-like display*

Trimension Workbench
- resolution of 1024x768
- FOV 120º
- ultrasonic tracking

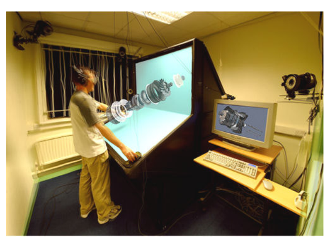


Figure 3-9: *Workbench display*

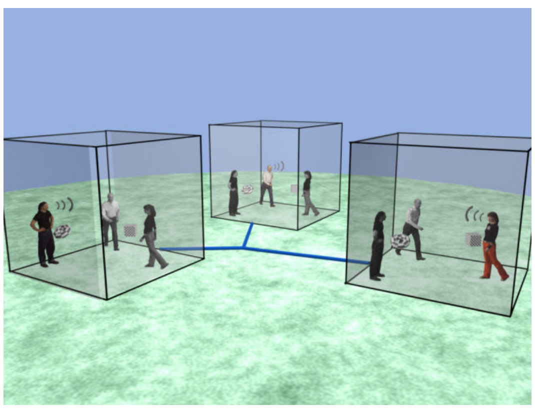
V8-HMD
- resolution of 640x480
- FOV 60º
- magnetic tracking

Figure 3-10: *HMD display*

In contrast to desktop systems, IPTs are not as limited in their support for non-verbal communication and naturalness of object interactions [Ståhl, 1999; Kjeldskov, 2001; Robinson et al., 2001]. The body movement is continuously tracked (usually just head and dominant hand), allowing both conscious and subconscious non-verbal communication to be captured and mapped onto the tracked person's avatar, as well as interaction with virtual objects. Current avatar solutions are not complete look-a-like representations of their user but technologies exist to create a close resemblance.

Because glasses have to be worn for stereo-vision, facial expression or eye gaze are difficult to track and therefore seldom found on avatars of immersive users. This creates less life-like collaboration compared to video images in video-conferencing or AR technology. However, as discussed in chapter 2 (see 2.2.3) postures are more important than gaze while there are interesting objects present.

Advances in immersive display devices are increasing their acceptance in industry and research [Brooks, 1999; Brodlie et al., 2004]. Their support of natural body and head movements may be used to view an object from every angle (Figure 3-11). An object can be reached for and manipulated with the outstretched hand, usually through holding an input device. The feeling of presence, and particularly the naturalness of interaction with objects, may be improved when the users can see their own body in the context of the virtual environment [Mine et al., 1997].



CAVE Gazebo [Roberts et al., 2003]

Figure 3-11: *Immersive Collaborative Virtual Environment*

The natural use of the body in IPTs to reference and interact with objects can increase both task performance and subjective impression of closely-coupled collaboration (see Chapter 6 for user trials) and an initial trial indicates that the scale of this improvement is relative to the spatial extent of the task [Roberts et al., 2005a]. Results indicate that communicative gaze plays a strong role in the performance of demonstrating focus of attention. As the physical extent of the display is considerably less than that of the shared environment, this advantage is unlikely to be connected to the mode of navigation. A more likely contributing factor is the reduction in fragmentation of the workflow, brought about by bringing two people within the same shared space and allowing each to see where the other is looking and pointing from a natural perspective.

Desktop CVEs use various methods to interact with objects in a virtual environment, such as go-go, ray-casting or occlusion techniques[2] [Poupyrev et al., 1998; Bowman et al., 2001]. These can be used in immersive CAVE-like displays, but have been originally developed using HMD. Desktop systems use 2D interface controls or virtual spheres or mouse picking, whereas immersive displays normally use one- or two-handed direct manipulation (virtual hand) using a tracking system. Evaluations of interaction techniques for immersive displays found that the virtual-hand is superior to ray-casting for the selection and manipulation of objects [Poupyrev et al., 1998; Steed & Parker, 2005]. A study by Byrne [1996] suggested that for some kinds of task, interaction is a more important facilitator of learning than immersion. Educational technologists reason that a student must interact with an environment for learning to occur [Anderson et al., 1995; Psotka, 1995]. However, the potential naturalness of interactions with objects in immersive VE makes interaction much easier and therefore more useful than in other types of remote environments.

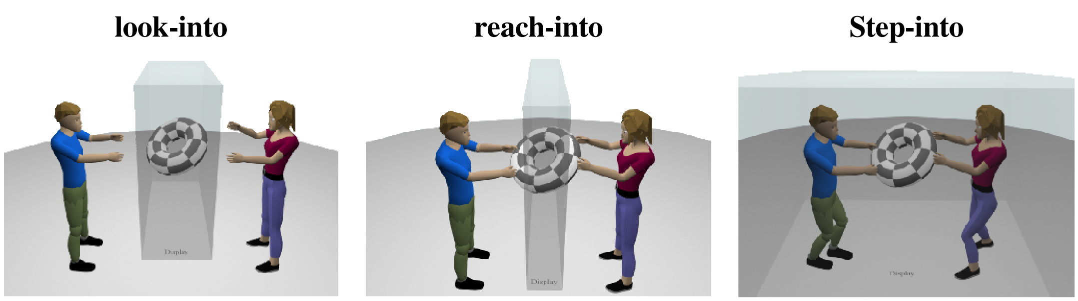
Comparisons of usability have been made between various immersive as well as desktop displays [Schroeder et al., 2001; Bowman et al., 2002b; Roberts et al., 2003; Sander, 2005] and they tend to show an advantage for immersion in various applications (see Chapter 5-7). This thesis is contributing to the research about immersive displays focussing mainly on aspects of closely-coupled collaboration and how various displays influence such collaboration. The next few Chapters (see Chapter 4-7) will introduce and discuss a number of user studies using immersive and non-immersive displays, whereas Chapter 8 will summarise various factors that can improve this collaboration. They will show that immersive displays are a useful tool to study user behaviour during distributed closely-coupled collaboration but also that a number of issues such as user interface or full body tracking could be improved to further enhance the experience of immersive users. First, however, the various technologies shall be summarised in their support for intuitive distributed object interaction.

3.6 Interaction Metaphors

Using a phone or text message to communicate can complicate collaboration due to possible misunderstandings arising from cues that cannot be communicated through the medium. The use of modern video-conferencing systems gives us more flexibility and

[2] go-go, ray-casting, occlusion techniques or others are various virtual arm extension to allow the manipulation of objects from a distance

support for non-verbal communication, such as facial expressions. Using video-conferencing, however, one only "looks into each other's world", which limits the operating range to move around shared objects and to be seen. In addition, shared object manipulation in current video conferencing systems is restricted to a window based 3D environment projected alongside the video window in which the remote participant is seen, for example AG-Juggler [Gonzalez, 2005]. The generic look-into metaphor is shown in Figure 3-12a and this is often extended, as is the case with AG-Juggler, by placing avatars in the shared space to represent the users. However, the movement of the avatars is controlled indirectly through mouse and keyboard and thus natural non-verbal communication is lost. In the case of Access Grid (e.g. [Childers et al., 2000]), the users are left to associate the video representations of remote participants in one window with the avatars in say an AG-Juggler window. In particular, it is hard to see how someone is interacting with an object when the operator, observer and object are each in separate windows, as in Access Grid.



(a) **(b)** **(c)**

Figure 3-12: *Interaction metaphors / participation frames*

Attempts are made to overcome the limitations of a "look-into" environment and to more closely reproduce a co-located setting. Such an alternative participation frame is "reaching-into each other's world" (Figure 3-12b). An example is the "Office of the future" where a user is sitting on a desk with a video screen attached to it [Raskar et al., 1998]. Thereby the user can interact with other co-located users, as well as the 3D video reconstruction of a remote collaborator. As users move, their locations are tracked so that the images are rendered from the correct perspective. The goal is for the remote room to be seen as an extension of the local room. In combination with augmented virtual objects (on the screen) it is possible for all participants to interact with these objects [Raskar et al., 1998; Baker et al., 2003; Yang et al., 2004; Barré et al., 2005]. The technology can therefore be considered as supporting a reach-into frame of participation. A constraint of this type of tele-immersion is that movements are restricted to the desk and it is still hard, compared to a real face-to-face meeting, to point to an object in the other's workspace unless the object is between the collaborators.

A possibility to avoid reference problems and other restrictions is to merge real and virtual world by using augmented reality technology. The users are no longer restricted by position and with co-located users share the same environment. However, a challenge is the problem of tracking and registration of real objects, people and the environment, in order for virtual models to be overlaid precisely on the real world.

Reaching-into someone's environment can be very beneficial for a number of tasks (e.g. [Ulhaas et al., 2001]) and augmented reality technology is a good example of how to realise this. But some tasks require more than just the representation of a few collaborators and few objects of interest. These tasks require the representation of a whole environment and the best way to interact with such an environment is to step-into it. For example, to train / simulate a rescue operation in a hazardous environment, the look and feel of this space is important and CVEs are a good technology to create such a space. More importantly, it even allows for creating environments impossible to (re-)create in the real world such as micro-spaces. This means that a system that allows users to share a common virtual space and to "step-into each others world" (Figure 3-12c), such as an immersive CVE, provides the closest resemblance to co-location. In a CVE, remote people and shared objects can be situated in a shared synthetic environment, in which one can navigate around and interact with a computer-generated representation of objects and other participants. Thus, whereas tele-conferencing systems allow people to look into each other's space, CVEs allow people and data to be situated in a shared spatial and social context.

3.7 Summary

In natural face-to-face collaboration, people use speech, gesture, gaze, and non-verbal cues to communicate. In many cases, the surrounding physical world and objects also play an important role, particularly in design and spatial collaboration tasks (Chapter 2). Real objects support collaboration through their appearance, physical affordances, such as size and weight, use as semantic representations, and ability to create reference frames for communication. In contrast, most interfaces for remote collaboration create an artificial separation between the real world and the shared digital task space. People looking at a projection screen or crowding around a desktop monitor are often less able to refer to objects or use natural communication behaviours. Observations of the use of large shared displays have found that simultaneous interaction rarely occurs due to the lack of software support and input devices for co-present collaboration [Pedersen et al., 1993].

Audio-only interfaces remove the visual cues vital for conversational turn taking, leading to increased interruptions and overlap, difficulty disambiguating between speakers and determining another's willingness to interact. With tele-conferencing, subtle user movements or gestures cannot be captured, there are few spatial cues among participants, the number of participants is limited by monitor resolution, and participants cannot readily make eye contact. Speakers also cannot know when people are paying attention to them or when it might be permissible to hold side conversations.

Table 3-1: *Support for remote collaboration and interaction*

	verbal	**non-verbal**	**shared objects**	**environment**	**naturalness**	**being-there**	**workflow**
	natural speech	*gestures, postures, facial expressions*	*artefacts of interest, person & non-person related*	*set the scene for natural collaboration and communication*	*intuitive performance of task*	*feeling of togetherness*	*continuation without technical interruptions*
face-to-face	natural	natural	shared / natural	shared by all	intuitive	high	synchronous and fluent
audio-conference	natural	NA	not shared	hear-into others	reduced	limited	interrupted through descriptions
groupware	NA	limited	asynchronous	shared	reduced	limited	asynchronous
tele-conference	natural	natural	not shared / natural	look-into others	intuitive	medium	continuous (if in frame)
tele-operation	natural	natural	semi-shared / naturalistic	look-into others	intuitive	medium	continuous (if in frame)
tele-immersion	natural	natural	shared / naturalistic	reach-into shared	intuitive	high	continuous (if in frame)
augmented reality	natural	naturalistic	shared / naturalistic	reach-into shared	naturalistic (with right tracking)	high	interrupted through visibility
typical CVEs	natural	unnatural	shared / unnatural	look-into shared	reduced	medium	interrupted through orientation
immersive CVEs	natural	naturalistic	shared / naturalistic	physical situated in shared	naturalistic (with right tracking)	high	continuous

In contrast, a number of challenges must be overcome before immersive or augmented reality technology is widely used for collaboration. Although shared interaction with objects is greatly supported, the capturing of nuances in body postures or gestures is depending on the technological effort invested. Further, gaze provides an important non-verbal cue in normal face-to-face and remote collaboration, yet current-generation displays or glasses cover the user's eyes.

A summary of the support for remote collaboration and interaction which is discussed in this chapter can be found in Table 3-1. It is possible to reduce the limitations and restrictions of computer mediation by enabling more flexible and natural interaction. Although the naturalness and intuitiveness of face-to-face communication is hard to achieve, immersive virtual environments provide additional and novel ways to enhance the weak areas of remote collaborative interaction.

Chapter 4

4 Evolution of a benchmark application - The Virtual Gazebo Prototype

The purpose of this research was to combine immersion and CVEs to study SHC (see Chapter2) and in particular closely-coupled collaboration. This chapter introduces a structured task of building a gazebo which was designed in order to examine distinct scenarios of closely-coupled collaboration and as a benchmark for further investigations. The following sections describe how it was tested and how findings lead to a rethink as well as redesign. Results of user trials are discussed in chapter 5-6 to keep the design process and experimentation of this benchmark application separated. The Virtual Gazebo design, implementation and testing were performed by two members of the research team, including this author. The detailed analysis of event handling and consistency control was done by Robin Wolff [Wolff, 2006], but some of the development inside is presented here in order to give the reader some understanding of the decision making during the application development and testing. But first a small introduction into CVEs shall be given introducing the basic principals and there support for closely-coupled collaboration.

4.1 Immersive displays connected via CVEs

Many team related tasks in the real world centre around the shared manipulation of objects. A group of geographically remote users can be brought into social proximity to interactively share virtual objects within a Collaborative Virtual Environment (CVE). CVEs are extensively used to support applications as diverse as medical treatment, military training, online games, and social meeting places [Pausch et al., 1992; Becker et al., 1998; Roberts, 2003; Prasolova-Førland, 2005].

Advances in immersive display devices are ensuring their acceptance in industry as well as research [Brooks, 1999]. Natural body and head movement may be used to view an object from every angle within an immersive display. The object may be reached for and manipulated with the outstretched hand, usually through holding some input device. The feeling of presence, and particularly the naturalness of interaction with objects, may be improved when the user can see its own body in the context of the virtual environment. Immersive Projection Technology (IPT) projects images onto one or more screens. CAVE-like IPT displays, such as a CAVE™ or ReaCTor™, surround the user with interactive stereo images, thus placing his body in a natural spatial context within the environment.

By linking CAVE-like displays through a CVE infrastructure, a user may be physically situated within a virtual scene representing a group of remote users congregated around a shared object. This allows each team member to use their body within the space to interact with other members of the team and virtual objects. The spoken word is supplemented by non-verbal communication in the form of pointing to, manipulating and interacting with the objects as well as turning to people, use of gestures and other forms of body language. This offers unprecedented naturalness of interaction and remote collaboration.

4.2 Principles of Distribution within CVEs

A key requirement of Virtual Reality (VR) is the responsiveness of the local system. Delays in representing a perspective change following a head movement are associated with disorientation and feelings of nausea [Lin et al., 2002; Meehan et al., 2003]. A CVE system supports a potentially unlimited reality across a number of resource-bounded computers interconnected by a network. The network, however, can induce perceivable delays in updating a distributed simulation [Roberts et al., 1996]. Key goals of a CVE are to maximise responsiveness and scalability while minimising latency. This is achieved through localisation and scaling [Roberts, 2003].

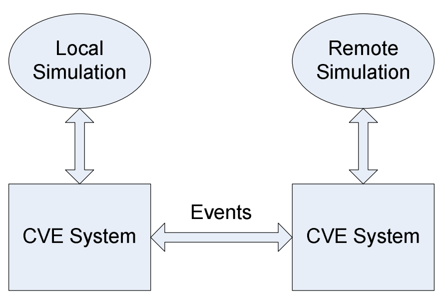


Figure 4-1: *Localisation through distribution*

Localisation is achieved through replicating the environment, including shared information objects and avatars, on each user's machine. Sharing experience requires that replications be kept consistent. This is achieved by sending changes across the network in the form of events (see Figure 4-1 for an illustration of localisation). Localisation goes further than simply replicating that state of the environment; it also includes the predictable behaviour of objects within it. Scaling limits the number and complexity of objects held on each machine and is generally driven by user interest [Greenhalgh, 1999]. The organisation and content of a Scenegraph[1] is optimised for the rendering of images. Although some systems [Park et al., 2000; Wolff et al., 2004] directly link scenegraph nodes across the network, most systems introduce a second object graph to deal with issues of distribution, known as the replicated object model. Objects contain state information and may link to corresponding objects within the local scenegraph.

A virtual environment is composed of objects, which may be brought to life through their behaviour and interaction. Some objects will be static and have no behaviour. Some will have behaviour driven from the real world, for example users. Alternatively, object behaviour may be procedurally defined in some computer program. In order to make a CVE attractive and productive to use it must support interaction that is sufficiently intuitive, reactive, responsive, detailed and consistent [Roberts, 2003]. By replicating object behaviour we reduce dependency on the network and therefore make better use of available bandwidth and increase responsiveness [Roberts, 2003]. Early systems replicated object states, but not their behaviour. Each state change to any object was sent across the network to every replica of that object.

4.3 Object interaction via CVEs

Most team work tasks require communication and many tasks, especially in design and science, require communication through and around shared objects. For example, a heavy table may require at least two people to carry it. The shared manipulation of objects requires consensus both at the human and system level. People need to agree where to carry an object and the CVE system needs to prevent network effects from producing confusingly divergent states of the object to each person. The latter issue was addressed for sequential sharing of objects within a ball game [Roberts et al., 1999]. The advanced ownership transfer allows instantaneous exchange of a ball between players in competitive scenarios. In IEEE 1516, concurrency control is defined to allow various attributes of a given object to be affected concurrently by distinct users. Sharkey et al. describe optimisations above the standard that allow control of an artefact to be passed to a remote user with little or no delay [Sharkey et al., 1998]. Elsewhere, a hierarchy of three concurrency control mechanisms is presented in Linebarger &

[1] Object oriented high level 3D graphics library

Kessler [2004] to tailor the problem of "surprising" changes during closely-coupled collaboration. A virtual tennis game is played between remote sites in Molet et al. and Basdogan et al. investigating the importance of haptic interfaces for collaborative tasks in virtual environments [Molet et al., 1999; Basdogan et al., 2000]. The authors stated that finding a general solution to supporting various collaborative haptic tasks over a network may be "too hard". A distinction is made between concurrent and sequential interaction with shared objects but this is not discussed further. As with Choi et al. [1997] a spring model is used to overcome network latencies to support concurrent manipulation of a shared object. Four classes of shared behaviour: autonomous behaviours, synchronised behaviours, independent interactions and shared interaction are introduced by Broll [1997]. Levels of cooperation within CVEs have been categorised by a number of research groups in similar ways: Margery et al. described the different levels of cooperation as level 1 - *co-existence and shared-perception*; level 2 - *individual modification of the scene*; and level 3 - *simultaneous interactions with an object* [Margery et al., 1999]. A similar taxonomy was presented for haptic collaboration that describes the respective levels as *static*, *collaborative* and *cooperative* [Buttolo et al., 1997]. Our studies provide a more detailed taxonomy of level 3, which will be described in section 4.4.

Much research has been dedicated to the development of CVE systems and toolkits. Some relevant examples are MASSIVE [Greenhalgh et al., 1995], CAVERNsoft [Johnson et al., 1998], DIVE [Frécon, 2004], NPSNET [Macedonia et al., 1994], PaRADE [Roberts et al., 1996] and VR Juggler [Bierbaum, 2001]. The DIVE system is an established testbed for experimentation of collaboration in virtual environments and, after three major revisions, remains an effective benchmark. The COVEN project [Frécon et al., 2001] undertook network trials of large scale collaborative applications run over the DIVE [Frécon, 2004] CVE infrastructure. This produced a detailed analysis of network induced behaviour in CVE applications [Greenhalgh et al., 2001]. DIVE was ported to CAVE-like display systems [Steed et al., 2001] and consequently an experiment on a non-coupled interaction task with two users in different CAVE-like displays was found to be very successful [Schroeder et al., 2001]. A stretcher application was implemented above DIVE, that investigated the carrying of a stretcher by allowing the material to follow the handles [Mortensen et al., 2002]. The work concludes that, although the Internet-2 has sufficient bandwidth and levels of latency to support joint manipulation of shared objects, the CVE did not adequately address the consistency issues arising from the network characteristics.

Several studies have investigated the effect of linking various combinations of display system on collaboration. It was found that immersed users naturally adopted dominant roles [Slater et al., 2000a]. A recent study by Schroeder et al. [2001], again using DIVE, investigated the effect of display type on collaboration of a distributed team. This work extended the concept of a Rubik's cube by splitting the composite cube such that two people could concurrently interact with individual component cubes while observing each other's actions. The study compared three conditions based on display combinations: two linked CAVE-like displays (symmetric combination); face-to-face; and a CAVE-like display linked to a desktop (asymmetric combination). An important finding was that the asymmetry between users of the different systems affects their collaboration and that the co-presence of one's partner increases the experience of the virtual environment (VE) as a place.

As DIVE is an established benchmark and to aid comparison to previous studies [Frécon et al., 2001; Greenhalgh et al., 2001; Schroeder et al., 2001; Steed et al., 2001; Mortensen et al., 2002; Frécon, 2004] it was adopted for this research.

4.4 The Virtual Gazebo Benchmark

The aim of this research was to study the impact of various levels of immersion on social human communication (SHC) and in particular closely-coupled collaboration. A benchmark was needed that would allow studying SHC during various scenarios of objects sharing (Table 1-2, p.5), while connecting various displays over a CVE. A further requirement was that it had to be easy enough to understand for students and other participants in the planed user trials as well as fast enough to be conducted (Figure 4-4). The choice was made to create a simple building task, the Virtual Gazebo.

Figure 4-2: *An ideal Gazebo*

Figure 4-3: *Tools used to build the Virtual Gazebo*

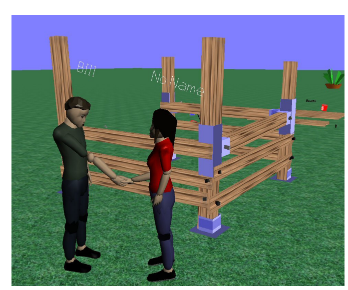


Figure 4-4: *Result of a 30min session*

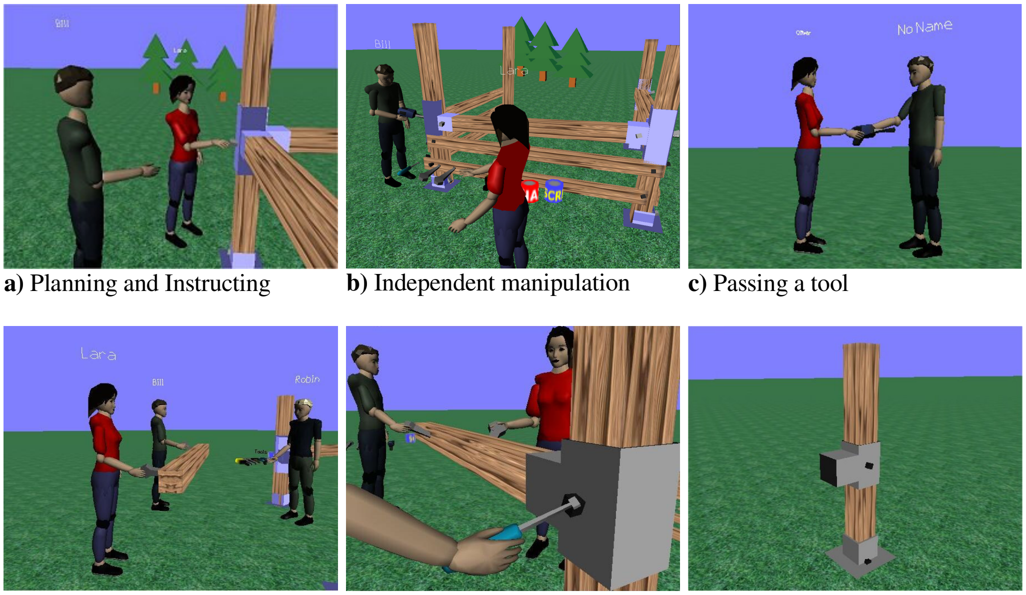
A gazebo is a simple structure that is often found at a vantage point or within a garden (Figure 4-2). A typical working environment contains materials, tools (Figure 4-3) and users. Wooden beams may be inserted in metal feet and united with metal joiners. Screws fix beams in place and planks may be nailed to beams. Tools are used to drill holes, tighten screws and hammer nails. To complete the Virtual Gazebo, tools and materials must be used in various scenarios of shared object manipulation, distinct in the method of sharing attributes. Scenarios include planning, passing, carrying and assembly (Table 4-1).

Table 4-1: *Forms of shared object manipulation with respect to timing*

Scenario &Timing	Fig. 4-5	Method of sharing	Example
planning synchronous	a	referencing objects and environment	discussion how to proceed
manipulating asynchronous	b	sequential manipulation of distinct object attributes	a person moves an object to a place, then another person fixes it
passing asynchronous	c	sequential manipulation of the same object attributes	a person moves an object to a place, then another person moves it further
moving synchronous	d	concurrent manipulation of the same object attributes	several people lift a heavy object together
assembling synchronous	e	concurrent manipulation of distinct object attributes	a person is holding an object while another person is painting it

These scenarios in Table 4-1 are a more detailed taxonomy of the three categorisation of cooperation from Margery et al. [1999]. They described three different levels of cooperation. In level 1 users can perceive and communicate with each other (see Figure 4-5a), while in level 2 they can individually modify the scene (see Figure 4-5b and Figure 4-5c). Within level 3, where the users can concurrently interact with the same object, we made a distinction between actions that are co-dependent (see Figure 4-5d) and those that have no direct effect on the others user action (see Figure 4-5e). The Virtual Gazebo study extended and clarified level 3 by the distinction between sequential and concurrent sharing of the same and different object attributes.

The Virtual Gazebo as a simple building task is easy enough to understand for novice users. However, a key requirement was the use of closely-coupled collaboration of two or more partners. In the real world this would be natural due to gravity and other constraints, in contrast, virtual environments could avoid these constraints making it difficult to study such collaboration. User would be able to finish the whole task on their own. To avoid this and to give the benchmark application a certain feeling of reality, constraints such as gravity, building order or the usage of tools were implemented as an integral part of the Virtual Gazebo.



a) Planning and Instructing **b)** Independent manipulation **c)** Passing a tool

d) Moving a Beam **e)** Fixing a beam **f)** A simple structure

Figure 4-5: *Collaboration scenarios when building a Virtual Gazebo (a-e)*

The implementation of the benchmark allows measuring application details (e.g. message transfer) to better understand network and consistency issues. These are in detail discussed in a parallel work [Wolff, 2006]. Measurements also allow the evaluation of task performance while questionnaires try to understand user perception. In addition, observations are used to gain further knowledge on the processes that are involved in closely-coupled collaboration between distributed places.

The next section will describe the various scenarios of shared object manipulation using an example within the Virtual Gazebo, after which successes and failure during the application development are discussed.

4.5 Building the Virtual Gazebo

On logging in, the user is placed in a garden, strewn with building materials and tools. Avatars appear, as the rest of the team enter the garden. “Wonder”-stacks keep the building site tidy by creating materials on demand. A user can take material from a nearby stack and use tools to fix it together to build the Virtual Gazebo. In the real world, constructing a garden gazebo on your own is not an easy task. To simulate the task similar to the real world some constraints were introduced. The simulation of gravity (object drops immediately to the ground) prohibits leaving materials in thin air and makes some materials too heavy to lift alone. The only task a single person can undertake is to drill holes and fit nails or screws. Moving, positioning and building all require teamwork. For example, one user must hold a joiner in place so that another user can fix it with a screw. In the following sections, five scenarios of planning, manipulating, passing, moving and fixing are examined, as summarised in Table 4-1, more in detail. The goal during many user trials was to build a simple structure (Figure 4-5f) that included all forms of object manipulation and allowed comparison between trials. Chapter 5-7 will describe in more detail experimentation details and display configuration used. For now, however, the next few sections describe the various interaction scenarios.

4.5.1 Planning and Instructing

The task of building the Virtual Gazebo routinely requires communication of the referencing of objects as well as the place within the environment that they are to be taken, (see Figure 4-5a). Communication of referencing must reflect nuances of speech and gesture and the interface must not restrict the recipient from capturing these. When using an IPT, control of gaze and pointing are driven through a tracking system and the user is surrounded by the display surface to the front, both sides and the floor. The complexity of the task requires the collaborative planning of a number of steps, which may involve several collaborators and objects. A wide field of view and direct control and communication of gaze and pointing should allow efficient referencing, location and identification of objects and other people.

4.5.2 Independent object manipulation

The task allows users to manipulate a number of objects independently without the input of a collaborator. For example, one user could fetch a tool and place it near the construction site while another user is positioning a foot-joint to lay the foundations of the Virtual Gazebo (see Figure 4-5b). The second user may then pick up a tool to drill a hole. This demonstrates the sequential manipulation of distinct object attributes. Users don't require communication for this scenario, yet the previous planning should secure fluent manipulation without fragmenting the workflow through too much independence.

4.5.3 Passing a tool

Various hand-held tools are fitted with the necessary attachments for construction. A drill makes holes in wood and metal, a screwdriver tightens screws and a hammer fixes nails. The garden only contains a single tool of each kind so that users will need to pass it between each other. Ideally, only one user can hold the tool at a time but can pass it smoothly to another user. Passing the tool (see Figure 4-5c), demonstrates sequential manipulation of the movement attribute as well as that of ownership. (With movement attribute it is referred to position and orientation, which may also be described by a path communicated between replicas.)

4.5.4 Moving a beam

When moving a beam, which is too heavy for a single user to lift alone due to simulated gravity, it becomes necessary for people to collaborate closely to synchronise their action in agreeing on the objective and to contribute in the same way to the task (see Figure 4-5d). This demonstrates concurrent manipulation of the position attribute as well as that of the orientation of the beam. Ideally, when two users attempt to drag the beam in opposing directions, it should move to a mean position between them. A pair of carry tools are used to pick up a beam, one at each end. When lifted by two carry tools, each end of the beam is attracted to its "connected" carry tool, as if by magnetism. This solution overcomes the issue of multiple parenting in the scenegraph and helps users to conceptualise the effects of network delays as magnetic attraction and inertia.

The users must synchronise their activity using any appropriate selection of forms of communication in order to move the beam to the desired position. It is up to the users to decide how they approach the task. They can use social communication and talk about how they proceed. In addition non-verbal communication can be used, such as gestures, to point where to go or which beam to take. When a user picks up one end of the beam with the carry tool, this end will be surrounded by a coloured aura, indicating to everybody that the user is now ready to drag the beam. The same happens when the second user picks up the other end. This is helpful for the synchronisation of the two users' actions.

4.5.5 Fixing a beam

Beams can be united with a metal joiner and screws. Again communication is important for coordination of the participants' actions. The users have to agree at which height to place the beam and of course on who is holding it in place. This decision process maybe influenced by the interface and it's capabilities to interact with the object. In addition, the user holding the beam needs to know when s/he can release it without dropping it to the ground. Every action with an object is indicated again by a coloured aura as a feedback of the progress of work. When the beam is fixed, the aura around the beam, joiner and screw disappears. Beams can be united with a metal joiner and screws. A joiner may be attached to a beam by drilling a hole through both and fixing with a screw. A second beam can then be fitted into that joiner in a similar manner. One person must hold a beam while it is attached to prevent it from falling (see Figure 4-5e). This demonstrates that one user is able to affect the attribute for fixing while another affects those of movement; in other words, the concurrent sharing of distinct attributes.

4.5.6 Finish the work

The various scenarios are continually applied by users of this benchmark application and not necessarily obvious as such. A multifaceted evaluation of various user trials combining questionnaires, task performance measurements and observations are described in Chapter 5-7. The evaluations will discuss the results and continue to describe how immersive CVEs might be used for distributed collaboration. First, however, a brief description about the benchmark development shall be given to better understand the complexity and the decision making process.

4.6 Application Design of the Virtual Gazebo

The Virtual Gazebo design, implementation and testing were performed by two members of the research team, including this author. The benchmark application was developed to work over the DIVE CVE [Carlsson & Hagsand, 1993]. DIVE was chosen because firstly, it is a well-established and widely accepted CVE platform, and secondly because of the ease of application development. The immersive extension Spelunk [Steed et al., 2001] allowed us to link various combinations of IPTs and desktop systems.

The user interface varied between used displays. Desktop systems had only mouse and keyboard for manipulation, interaction and locomotion, whereas IPTs had a tracked wand including a joystick available. The latter allowed simple gesturing and grasping an object was carried out by moving to an object (physically within the IPT or using the joystick) and bending the body such that the virtual hand intersected with the object. Now pressing a button on the tracked wand allowed free movement and rotation of this object. In contrast, the desktop user had to select an object using the mouse cursor and pressing a key at the same time. Various key-mouse combinations allowed manipulation and rotation in any direction.

The application is implemented as a set of interactive objects. Three classes of objects were needed: avatars, materials and tools. Each object has a graphical representation and all have scripted behaviours that match their purpose. In DIVE, all objects are structured hierarchically in a distributed database. Their current state is represented by attributes, which may be brought to life by user-defined object behaviour scripts. All behaviour scripts are reactive and triggered by specific DIVE events. These are update messages, generated by the CVE system to update replicated versions of the distributed virtual environment. DIVE supports several event types. These include object transformation events, such as movement or rotation; object interaction events, such as grasp, release or select events; object collisions; and changes to object-specific properties and flags. Most functionality of the Virtual Gazebo is triggered by collisions of material and tool objects. For example, when a drill tool is held closely to a material object so that they collide, the resulting collision event, generated by the system, would trigger a procedure in the material object's behaviour script to increment a "hole-counter" property. A detailed description of the objects' behaviour and other implementation details can be found in Roberts et al. [2005b].

4.7 Initial Experimentation and Results

The first prototype Virtual Gazebo application has been tested between IPTs and desktop located at the University of Reading (UK) and University College London (UK). Spelunk, an immersive extension to the DIVE CVE [Steed et al., 2001], was used to link the IPTs. Here, findings are presented of the first application prototype which was regularly tested between sites over a three weeks period. The goal during many user trials was to build a simple structure (Figure 4-5f) that included all forms of object manipulation and allowed comparison between trials. Observations where taken about the CVE performance, application behaviour and user actions.

DIVE uses multicast, which is used extensively by many CVE systems to increase scalability of group communication. Although multicast works within a local area network, it is usually necessary to tunnel multicast packets between local area networks, particularly when they are separated across the Internet. DIVE proxy servers [Frécon et al., 1999] were used to tunnel packets between local area networks at each site. Audio communication was supported through the UCL Robust Audio Tool (RAT) [Hardman et al., 1995]. Desktop users where not provided with a headset for audio communication, but a speaker and microphone near the monitor. This was done to hide some of the technology, only the IPT user had an attached microphone.

Using the first prototype of the Virtual Gazebo, each user was able to interact with objects successfully and it was generally easy to interpret what remote users where doing, especially with the support of audio communication. The actions and gestures of tracked users were much easier to understand than those of desktop counterparts.

Three problems, however, severely hampered collaboration around shared objects: Firstly, the (loose) ownership mechanism in DIVE made it difficult for two users to carry a beam concurrently. Secondly, many important interactions with shared objects were not being reflected remotely, such as creating objects or grasping tools. Thirdly, one would have expected verbal communication between remote users to become more natural when the technology is transparent, that is when the microphone and speakers are hidden. However, the opposite was the case, users did rarely make use of audio communication which had a negative effect on the collaboration.

With these problems it was very difficult to build the Virtual Gazebo. The beam was lightened so that one user could lift it and test trials were repeated to see what could be achieved. Users in a link-up between the three IPTs created what resembled a sloppily constructed corral or sheep pen. In addition users experienced communication problems, as it was not always clear if an action such as drilling a hole was successful due to the lack of application feedback.

Extensive tests were undertaken to verify a hypothesis of event loss due to the high tracking update rate of the immersive display and why this should be a particular problem for shared manipulation between IPTs [Wolff et al., 2004]. The movement of avatars, materials and tools all increased during shared manipulations, causing bursts of events at exactly the time when reliability and low latency were needed. These bursts were evident in latencies rising to several seconds for scenarios such as fixing beams with a joiner. It was found that the problem did not arise when representing a desktop user interacting with an object. The avatar used to represent a desktop user is simpler than that used for the user of an IPT. A simpler avatar was tried to represent the immersed display user and this was found to solve the problem. The new avatar had no moving parts and thus produced less network traffic to update. Although this avatar solved one problem, its simplicity made human-like, non-verbal communication much harder.

DIVE incorporates an optional reliable message service, Scalable Reliable Multicast (SRM) [Floyd et al., 1995]. When enabled, SRM ensures all messages from that user's device are delivered. Enabling SRM, while using the more complex avatar, ensured the representation of the effect of remote interactions with an object. The drawback of using SRM was a lag of greater than a second in the representation of the actions of a remote user, including movement and interaction. Such a delay makes successful communication and interaction very difficult.

4.8 Improved Application Design of the Virtual Gazebo

The earlier trials showed that the implementation of the Virtual Gazebo prototype application lacks heterogeneous mechanisms for concurrent sharing of objects and that the tracking of the immersive display was responsible for a high amount of event occurrences resulting in event losses of important messages, such as fixing an object.

To overcome the problem of the high amount of event occurrences, introduced by the tracking system, the used DIVE version was enhanced with an event filter. This filter was setup to send update events of user movements only if this movement was above 1cm. It was found that this is a good compromise between sufficient detail to support understandable non-verbal communication and sufficient synchronisation to achieve shared object manipulation. Many events below this seemed to be caused by tracking jitter rather than real user movement. This reduced the frequency of events and allowed us to use our more human-like avatar.

For supporting the synchronisation of interactions and communication, a feedback in form of visual clues was implemented. This feedback was a colour-change in a number of objects, for example, adding an additional transparent coloured aura around one end of a beam as soon as it has collided with a carry tool. Roberts et al. [2005b] describe these and other changes in more detail. Finally, headsets for audio communication were again provided to ensure that users can hear each other and that the microphone will pick up the voice.

4.9 Repeated Experimentation and Results

Again it was attempted to build the Virtual Gazebo in a number of linkups between the various sites with tuned event communication to gain acceptable levels of latency and reliability. Figure 4-4 (p.61) shows the result of a successful collaboration of about 30 minutes.

In order to obtain a workable level of reliability, while three immersive users shared the manipulation of objects, it was necessary to reduce the rate of sending of avatar movement to the network to 5Hz while maintaining 10Hz for the shared objects. For example, this was found sufficient when three immersive users fixed a beam, one holding the beam, another drilling the hole and a third inserting a screw. In order to gauge latency between displays users undertook what we called a wave test. A user at UCL moved his hand up, down, left then right, speaking the movements as he did them. At a 5Hz update rate, these movements were reflected in Reading before the spoken word, suggesting that the CVE had less latency than the audio tool. Network latencies between Reading and UCL typically varied between 15 and 25ms. The reduced update rate of avatars resulted in less natural movement making it slightly harder to interpret their actions.

Objects still became un-graspable after they had been picked by another user but far less frequently than in the original Virtual Gazebo. The reliability of infrequent state changes such as drilling holes and inserting screws was also increased. Infrequent loss of such changes was often overcome through teamwork. For example, if the creation of a hole is not reproduced at all sides, users can report this and ask for another hole to be drilled. In a later stage of the research, the application was implemented by the author as a C-plugin for DIVE and no longer as Tcl/Tk scripts. This seemed to solve the issue of losing events with the drawback that application worked only as long as the first machine running the plugin was online. By then a new CVE [Wolff, 2005; Wolff, 2006] was under development and further investigation into this issue was put on hold.

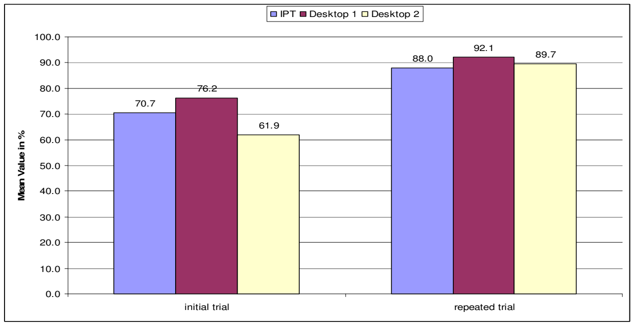


Figure 4-6: *Perceived verbal communication usage for the initial(section 4.7) and revised (section 4.9) prototype of the Virtual Gazebo*

The introduction of carrying tools enabled joint manipulation of beams. Human communication and visual clues helped to synchronise lifting of the beam and choosing a direction in which to carry it. Although latency was not apparent in avatar movement, remote manipulation of the beam was often delayed by up to one second. This resulted in wild beam movement not unlike that of a rodeo horse. The fact that only one object was affected suggests a backlog of interpreted script events as opposed to filling of a receive buffer, which would have affected all objects. A change of the tracker rate limit from 100ms to 200ms seemed to solve the delay problem.

The introduction of feedback in form of visual clues after an object interaction was found to be very successful. Observations and personal experiences showed that visual feedback allows judging the progress and if necessary let other collaborators know about the repetition of action. For example, a user sees an object fixed while his partner may not (due to consistency issues), but because the partner did not see a visual clue he can ask for a repeat of fixing. In addition, in a trial with one immersive and two non-immersive, the induction of headsets was found to increase verbal communication (see Figure 4-6). This was found after statistical analysis with ANOVA[2] to be significant ($F(1,96)=26.71$, $MSW=1.79$, $p=0.000$) and observations showed that the teams worked together more successfully.

4.10 Summary

The development of the Virtual Gazebo benchmark application has demonstrated that users, sharing the manipulation of objects, can adapt to the limited effects of remoteness between networked CAVE-like displays. Limiting these effects, however, required considerable effort in application development and deployment. Although

[2] see Chapter 5 for introduction about ANOVA analysis used in the user trials

many CVEs provide mechanisms for dealing with the effects of remoteness, these are barely sufficient for such linkups and require a combination of application constraints and workarounds as well as fine tuning of event communication. CVEs have been routinely used for linking desktop display systems over a decade, but the use of immersive displays allows better communication and interaction while introducing new challenges. This concurs with earlier work by Slater et al. [2000a] that it is easier to collaborate with a remote user when their avatar is driven by tracking data. IPTs are different because the users are tracked and the communication of tracked human movement is data-intensive. This problem is exacerbated by the very different data requirements of shared object manipulation, where occasional vital events must be sent reliably and in order, often coincident with bursts of non-vital movement events.

In addition, one would have expected verbal communication between remote users to become more natural when the technology is transparent, that is when the microphone and speakers are hidden. However, a significant increase in verbal communication could be observed when the user is constantly aware of a familiar communication device, that is, a headset with microphone and earphones. When this was introduced, the team seamed to work together more successfully and were more engaged in the task.

Chapter 5

5 Testing the benchmark in immersive and desktop displays

This chapter describes employment and user trials of the benchmark application introduced in the previous chapter. The benchmark has been used over the last few years by this research team to study human interaction and more specifically closely-coupled collaboration in CVEs. The structure of this and the following chapters follows the development and research timeline of the last few years, in order to demonstrate the research process of how and why certain user trials where undertaken.

5.1 Hypothesis-1: Immersive displays (CAVEs) are suited for closely-coupled collaboration

A study by Broll in the mid-nineties, concluded that concurrent shared manipulation of objects in a CVE would not be possible with technology at that time, due to delays caused by distribution [Broll, 1995]. At the beginning of this research in late 2001 the problem was still not addressed, although some research was trying to investigate the possibilities of distributed collaboration, at first using single desktop system (e.g. [Ruddle et al., 2002]) and later through networked immersive displays (e.g.[Mortensen et al., 2002; Linebarger et al., 2003]). Therefore, in order to examine distinct scenarios of sharing the manipulation of an object, a benchmark application was developed and put on trial. The benchmark, the Virtual Gazebo, and its development is described in detail in the previous Chapter 4. Section 5.2-5.4 introduces the task and the setup for the various displays. The results are given in section 5.5 & 5.6, discussed throughout in relation to previous studies and finally summarising in section 5.7.

5.2 Task breakdown

Variations of the Virtual Gazebo have been built during several collaborative sessions involving desktops and CAVE-like displays at Reading and London in the UK, (Figure 4-4, pp. 61). As in the real world, building a gazebo can take several hours of often repetitive work. Thus, for detailed evaluation the task was reduced to construct a simpler structure (Figure 5-1), removing unnecessary repetition but still requiring all forms of object sharing (Table 4-1, pp. 62) along with varied human communication. The detailed breakdown of the new task is given in Table 5-1, where an example is shown of how two users might construct the simple structure. A third user may assist by, for example, fetching a tool while two others are carrying the beam or may help with task planning and execution.

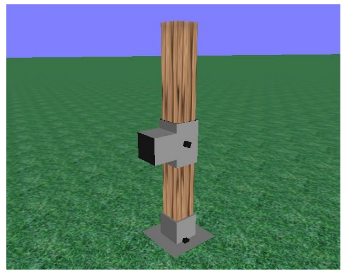


Figure 5-1: *The task target of a simple structure*

Table 5-1: *Detailed task breakdown showing example collaboration*

Sub-task	**Description**	**User 1**	**User 2**
ST1	Place foot	Fetch foot and place squarely on the ground	
ST2	Carry beam	Fetch carry tools and use one to lift each end of the beam. When both ends are lifted, carry the beam to the foot	
ST3	Place beam in foot	Place one end of the beam in the foot	And then lift the other end so that the beam is vertical
ST4	Drill hole	Fetch the drill and drill a hole through foot and beam	Hold the beam in place
ST5	Insert screw	Fetch a screw, insert it in hole	Hold the beam in place
ST6	Tighten screw	Fetch a screw driver and tighten screw	Hold beam in place until screw tightened
ST7	Place T joiner	Fetch T joiner and hold it in place on the upright beam	
ST8	Drill hole	Hold the T joiner in place	Fetch drill and drill a hole through foot and T joiner
ST9	Insert screw	Hold the T joiner in place	Fetch a screw and insert it in the hole
ST10	Tighten screw	Hold the T joiner in place until screw tightened	Fetch screw driver and tighten screw

The methodology for evaluating the task is explained both for team performance (5.5) and subject perception (5.6). Team performance measures the time taken to complete the task and each component sub-task. User evaluation details the responses to a questionnaire on the perception of the collaboration.

5.3 Display configurations

The tests involved distinct display configurations, all different in their ability to facilitate interaction with the other two participants as shown in Table 5-2. Two basic display types were used, a CAVE-like cubic IPT, and a desktop. All of the immersive configurations restricted the user to one-handed interaction within our application. Collaboration would still have been necessary for two-handed input because of the effect of gravity on "heavy" beam objects.

Table 5-2: *Display configurations*

Name	**IPT1**	**IPT2**	**DT1**	**DT2**
Location	Reading	UCL	Reading	Reading
Display	4 wall cubic display	4 wall cubic display	Desktop	Desktop
Input	Tracked Wand incl. a Joystick	Tracked Wand incl. a Joystick	Keyboard & Mouse	Keyboard & Mouse
Tracking	Ultrasonic/Acoustic Intersense IS900	Ultrasonic/Acoustic Intersense IS900	none	none
Computer	2 pipes 6 dedicated processors SGI Origin 2000	2 pipes 12 dedicated processors SGI Origin 2000	2.4 GHz Dell PC with Nvida Quadro 960	2.4 GHz Dell PC with Nvida Quadro 960
Audio	Yes	Yes	Yes	No*
Embodiment	Motion tracking	Motion tracking	Low realism	Medium realism
Tracking filter	1cm / 200ms	1cm / 200ms	none	none

* within calling distance to DT1 (approx. 2m) and able to use text chat

Tests were undertaken over a six month period. Typical network latencies during this period were:

- Reading to London: 19ms
- Reading to Reading: 17ms (through slow switch to simulate national Internet latency)

The DIVE CVE was used for experimentation as described in the previous Chapter. DIVE version 3.3.5 was used on all devices, which was extended with an event monitoring plugin and an event filter. The event monitor timed event callbacks with synchronised clocks [Anthes, 2002]. Event filtering reduced the frequency of events generated by the tracking system. Throughout the tests, the tracking system was filtered to only produce events for movements greater than one centimetre. In extensive testing (see Chapter 4), this level of filtering was found to produce the optimal balance between system performance and usability.

5.4 Test Conditions and Questionnaire

Two user trials where undertaken to compare desktop and CAVE-like displays. The initial first trial compared team performance between novice and expert users while the second trial was testing the Virtual Gazebo application with a large number of student volunteers. All participants signed a consent form (see Appendix A) and novice users had never previously used the Virtual Gazebo. At the beginning of each user trial participants were briefly introduced into the interface of each display and application.

In an initial trial multiple test-runs compared the performance of both expert and novice teams across the display configurations IPT1, DT1 and DT2 (see Table 5-2). A detailed description of the trial setup is given below in section 5.5. In the second trial 48 student volunteers were split into teams of three for each test. They were told to build a structure similar to Figure 5-1 and that this would require their collaboration. Participants were advised by the investigation team to use the provided communication resources and to help each other if possible. The perceived effectiveness of collaboration involving shared objects and the perceived effect of display type were investigated, using a user evaluation questionnaire. Within every trial, each user interacted through a distinct display device and was questioned on his perception of the effectiveness of teamwork. The conditions C1-C3 of Table 5-3 shows how the three users were assigned to the displays as this defined the user references inside the questionnaire. For example, condition C1 questioned how the user of IPT1 perceived the effectiveness of collaboration with the users of desktop DT1 and DT2. In condition C4 an additional ten teams were used of which two local users (at University of Reading) were assigned to IPT1 and DT1 while a third user was connected from the IPT2 at University College London (UCL).

Table 5-3: *Test conditions*

Condition	Questioned user	User 2	User 3
C1	IPT1	DT1	DT2
C2	DT1	IPT1	DT2
C3	DT2	IPT1	DT1
C4	IPT1	IPT2	DT1

The questionnaire was aimed at ascertaining the user's subjective perception of collaboration, both generally and for each specific task. Questions were based on those of Usoh and colleagues [2000]. Answers could be given on a Likert-type scale [Sitzman, 2003] of 1-7, where 1 represented agreement to a very small extent and 7 to a very large extent. Those scales were later converted to percentage in order to allow for better and clearer comparison. Errors arising from a user's misinterpretation of a question were reduced by asking sets of related questions. For example, "to what extent did the two of you collaborate" was contrasted with "to what extent did each user hinder the task". During the analysis where differences were observed for those alternating questions, a further analysis is performed and documented below. Summary findings of the questionnaire (see Appendix A.1) are represented in the next sections of this chapter. Thereby a special accent is given to questions related to social human communication and the shared manipulation of objects.

5.5 Team Performance

Team performance was measured both in terms of time taken to complete the task and each component sub-task [Roberts et al., 2003], in order to gauge the support for collaboration offered by various display configurations. Multiple test-runs compared the performance of both expert and novice teams across the display configurations IPT1, DT1 and DT2. Timing of novice users came from user trials with 12 student volunteers while additional trials looked at the effect of display combination on expert users. The experts had used the Virtual Gazebo over several months. The teams were left to determine their own organisation of roles in a natural way as the task progressed. The only constraint was the order of the sub-tasks ST1 to ST10 as described in Table 5-1.

5.5.1 Collaboration between novice subjects

The trials were started with a set of 12 novice users, each of whom undertook the trials voluntarily and all were students of undergraduate programmes at the University of Reading. None had previous experience of working in an immersive display or of the Virtual Gazebo application. Teams of three subjects performed the task in three test-runs using IPT1, DT1, DT2. All of these display systems were local at the University of Reading. By changing places between test-runs, each subject interacted through the entire set of display configurations in the same geographical location (C1-C3 of Table 5-3).

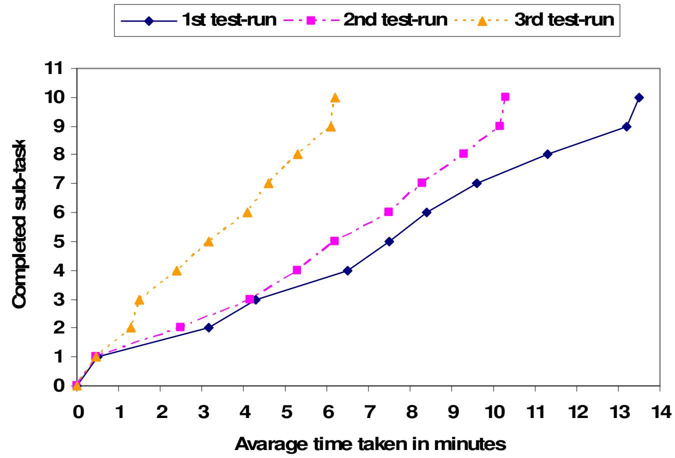


Figure 5-2: *Timing example of a novice team for three tests-runs*

Figure 5-2 shows the measured timing of team of novice users for each completed sub-task. For all four novice teams a strong correlation was observed between the experience of users and the time taken to complete the task. Subjects that faced our test environment for the first time appeared to have difficulties in recognising the constraints of the application and the handling of the interface. However, both were learnt quickly resulting in a doubling of performance by the third attempt (e.g. after performing first C1, next C2 and last C3).

5.5.2 *Effect of display configuration on expert users*

Performance measurements for novice subjects vary greatly. Consequently, to better gauge the effect of display combinations, repeated test-runs between pairs of expert subjects were undertaken. The set of expert subjects had several months of regular experience of both the Virtual Gazebo application and the interface. Firstly, display configurations were compared as before (TRA-TRB) and then this was repeated with constraining subject roles (TRC-TRD). The latter was done to gain a clearer understanding of the effect of role on subject performance for a given display. The constrained roles were divided into primary and supporting, the former undertaking the more difficult parts of subtasks, such as fixing, while the latter held material in place. This could only have been done with expert users as they had extensive exposure to application and interface, and it would have been difficult to explain to novice users to constrain their activity. The idea of this expert test was to gain further knowledge on role-taking. Table 5-4 distinguishes the test-runs undertaken by expert teams. In contrast, in the first test-run TRA roles were not forced to primary or supporting.

Table 5-4: *Overview of roles in expert users test-runs*

Test-run	**IPT1**	**DT1**	**DT2**
TRA	Unconstrained	Unconstrained	-
TRB	-	Unconstrained	Unconstrained
TRC	Primary	Supporting	-
TRD	Supporting	Primary	-

Figure 5-3: *Timing of a expert team for four different test-runs*

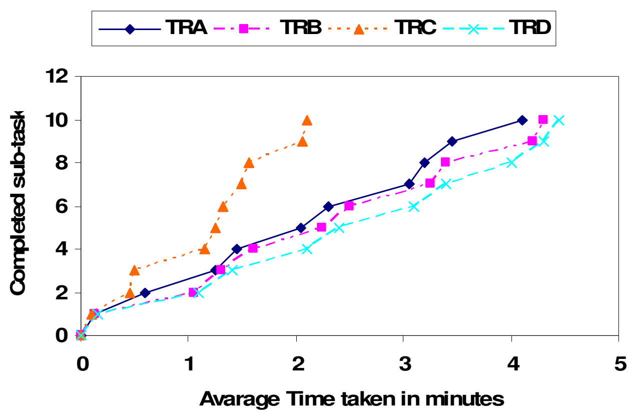


Table 5-5: *Performance increase if IPT has leading role and DT supporting role*

session	*example (Predominant activity)*	*increase TRC→TRD*
	Start	
ST1	Place foot (moving)	48 %
ST2	Carry beam (moving)	35 %
ST3	Place beam (positioning)	73 %
ST4	drill hole (use tool)	44 %
ST5	Insert screw (positioning)	53 %
ST6	fix beam (use tool)	65 %
ST7	Place T joiner (positioning)	64 %
ST8	drill hole (use tool)	55 %
ST9	Insert screw (positioning)	65 %
ST10	fix T joiner (use tool)	65 %

Figure 5-3 shows the timing of the team of expert users. For unconstrained roles (TRA), the expert teams took about half of the average time taken by novice teams. In addition, graphs TRA and TRB in Figure 5-3, reveal that the type of display does not make a clear difference when the organisation of role is unconstrained. However, giving the primary role to the CAVE-like display user, results in a considerable performance increase, see TRC versus TRD in Table 5-5.

The taking of the primary role by the immersed user results in a clear performance increase for most subtasks. The advantage appears to relate more to the suitability of each interface to a given form of object manipulation, rather than to the method of object sharing. Over the whole task, a cumulative performance increase of 55% was measured for the CAVE-like display against the desktop.

Observation also showed a performance difference between immersive and desktop users. The latter mainly had difficulties with the complicated keyboard shortcuts to control various aspects of the application such as avatar and object rotation as well as translation. The greater perception of depth in immersive displays was also helping to grasp objects with ease and then position them. Further observations and scientific analysis of user perception is the focus of discussion in the next section of this chapter.

5.6 User Evaluation

A large scale trial (for statistical analysis) with over fifty volunteers was undertaken to evaluate user perception of novice teams in various display combinations (see general description of Test Conditions and Questionnaire above in section 5.4). For the analysis of the questionnaire the statistical approach of analysis of variance (ANOVA) was used to verify the significance of the results. It is used to pinpoint the sources of variation from one or more possible factors. ANOVA helps determining whether the variations are due to variability between or within methods. The within method variations are due to individual differences within trial groups, while the between method variations are due to differences between the methods. A result of ANOVA is to test for differences among all the population means with a significance deviance α. The limit in this case used for significant deviance was α=0.05 (95% confidence interval). The results are given with MS_W as the mean square within groups, F(a,b) as the variance between groups/MS_W and *p-value* as the actual deviance, with four decimal places. A posthoc Tukey test was applied if a significant difference could be found to clarify between which groups those differences appeared.

The questionnaire was divided into different sections trying to examine various aspects of the user trial. The first two sections aimed to understand how participants perceived their performance and collaboration between the two distinct tasks of concurrently sharing the same (carrying) and different (fixing) object attributes. The last set of questions aimed to identify the social and human correlations. The answers to each question were collected and passed through an ANOVA analysis, where display configuration was used as factor (in Minitab[1] statistical package).

[1] Minitab release 13.20 from Minitab Inc.

5.6.1 Contributions to carrying a beam

A first question to evaluate the concurrent sharing of the same object attribute (carry beam) was "*To what extent did each person contribute to the task while carrying a beam?*" an analysis of variance (ANOVA) demonstrated that there was a significant difference between the conditions, see Table 5-6. Conditions C1, C2 and C4 all showed a clear statistical significance, while C3 showed a close statistical significance. An ANOVA across the combined questions, for conditions one to three, illustrates that there is no significant difference between all answers ($F(2,41)= 0.18$, $MS_W=14.9$, $p=0.840$) and thus the answers may be united, across conditions C1 to C3, to gain a better statistical certainty of device importance. The ANOVA for this demonstrated that there is a significant difference and a posthoc test showed that the difference lies between all three devices. These results show that asymmetry in linked devices (CAVE-like vs. desktop) affects perceived contribution. Immersive users are considered by all to contribute more than desktop users. Furthermore, where a team comprised of two immersed and one desktop user, the latter was left out of most of the activity. The significance of this finding is demonstrated through the ANOVA of C4 that had a statistically significant *p-value* (0.003).

Table 5-6: *ANOVA results for contribution to carry a beam*

Condition (%)	IPT1		DT1		DT2		IPT2		ANOVA results (α=0.05)
	mean	SD	mean	SD	mean	SD	mean	SD	
C1 IPT1-DT1-DT2	81.0	17.7	67.5	23.9	54.3	29.7	-	-	$F(2,48)= 5.12$, $MS_W=2.79$, $p=0.010$
C2 DT1-IPT1-DT2	83.5	20.9	65.5	28.2	53.4	27.5	-	-	$F(2,34)= 4.67$, $MS_W=3.21$, $p=0.016$
C3 DT2-IPT1-DT1	77.9	25.0	65.5	23.2	51.4	31.0	-	-	$F(2,30)= 2.65$, $MS_W=3.40$, $p=0.087$
C1-C3	81.0	20.4	66.3	24.4	52.9	28.5	-	-	$F(2,118)= 12.96$, $MS_W=2.94$, $p=0.000$
C4 IPT1-IPT2-DT1	67.9	29.3	31.0	10.8	-	-	78.6	20.2	$F(2,19)= 8.29$, $MS_W=2.44$, $p=0.003$

Where: α is the limit of significant deviance
MS_W is the mean square within groups
$F(a,b)$ is the variance between groups / MS_W
p is the actual deviance, with four decimal places
M is mean & SD is standard deviation

significant difference between: [green] and [orange]
as verified by the posthoc test (Tukey)

5.6.2 Contributions to fixing a beam

The same question was asked for the task of concurrent sharing of different object attributes (fixing a beam) and as Table 5-7 shows no significant difference for the first three conditions could be found. An ANOVA across the combined questions of the conditions one to three, showed that there is no significant difference between all answers. The ANOVA for the united questions C1-C3 demonstrated a close to significance *p-value* (0.097) and a posthoc test showed that the difference lies between IPT1 and DT2. These results show that the effect of asymmetric devices is perceived to play considerably less of a role in the level of contribution, in fixing a beam than in carrying it. The actual deviance for fixing is 0.097 compared to zero for carrying. However, for C4 similar significant findings where found, compared to the previous analysis of concurrent sharing of the same attribute.

Table 5-7: *ANOVA results for contribution to fix a beam*

Condition (%)	IPT1		DT1		DT2		IPT2		ANOVA results (α=0.05)
	mean	SD	mean	SD	mean	SD	mean	SD	
C1 IPT1-DT1-DT2	78.2	17.6	64.7	25.4	63.9	25.8	-	-	$F(2,50)= 2.05$, $MS_W=2.66$, $p=0.140$
C2 DT1-IPT1-DT2	77.4	18.7	73.8	24.2	64.3	27.6	-	-	$F(2,33)= 0.97$, $MS_W=2.78$, $p=0.389$
C3 DT2-IPT1-DT1	64.9	17.3	67.9	15.1	61.9	26.8	-	-	$F(2,32)= 0.25$, $MS_W=2.05$, $p=0.777$
C1-C3	74.3	18.4	68.1	22.4	63.4	26.0	-	-	$F(2,121)= 2.38$, $MS_W=2.48$, $p=0.097$
C4 IPT1-IPT2-DT1	76.2	14.8	32.1	21.4	-	-	54.8	30.5	$F(2,13)= 4.30$, $MS_W=2.69$, $p=0.037$

Where: α is the limit of significant deviance
MS_W is the mean square within groups
F(a,b) is the variance between groups / MS_W
p is the actual deviance, with four decimal places
M is mean & SD is standard deviation

significant difference between: [green] and [orange] as verified by the posthoc test (Tukey)

5.6.3 Comparison of perceived contribution for carrying and fixing

The difference of the effect of asymmetric devices observed when carrying as opposed to fixing the beam is confirmed in Figure 5-4, which combines the above results.

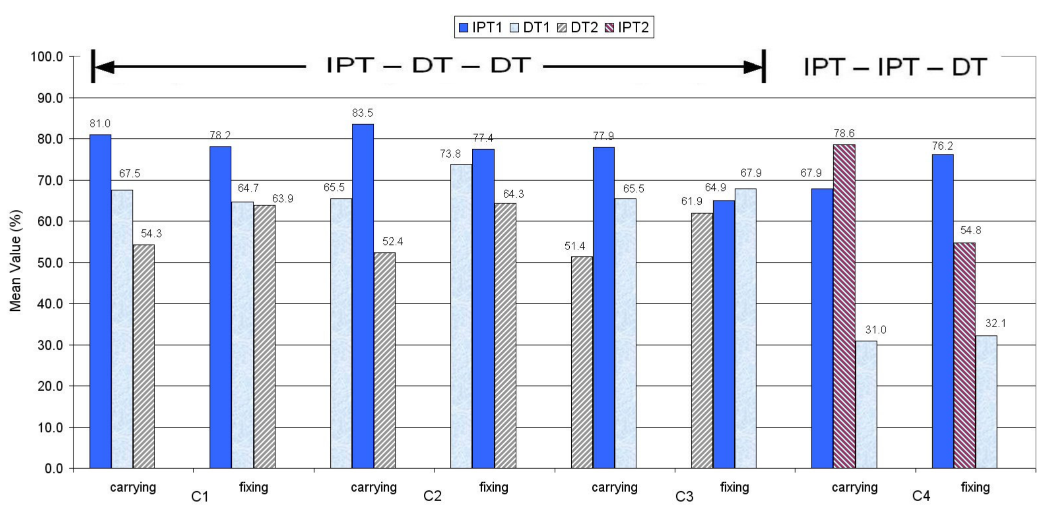


Figure 5-4: *Perceived contribution while both carrying and fixing the beam*

5.6.4 *User hindrance of the task*

In answer to the question "*To what extent did each user hinder the task?*" an ANOVA unveiled that there is no significant difference between the conditions, p=0.699 for carrying a beam and p=0.846 for fixing a beam. The results for carrying a beam M=44.9, SD=23.4 and for fixing a beam M=46.2, SD=21.1 indicates clearly that the participants did not excessively hinder each other.

5.6.5 *Collaboration between users*

Carrying and fixing a beam requires collaboration between two users. When it comes to the evaluation of "*To what extent did the two of you collaborate?*" and "*How well did you and the other person together performed the task?*" an ANOVA showed only a significant difference (p=0.002) in C4 for carrying the beam (M=80.4, SD=25.3), while there was no significant difference in any of the other trials, neither for carrying nor for fixing a beam, Figure 5-5. These results show that from the perspective of immersed users, collaboration is considerably easier with a symmetric user (another immersive user). However, desktop users found the type of remote display to play little part in the level of collaboration.

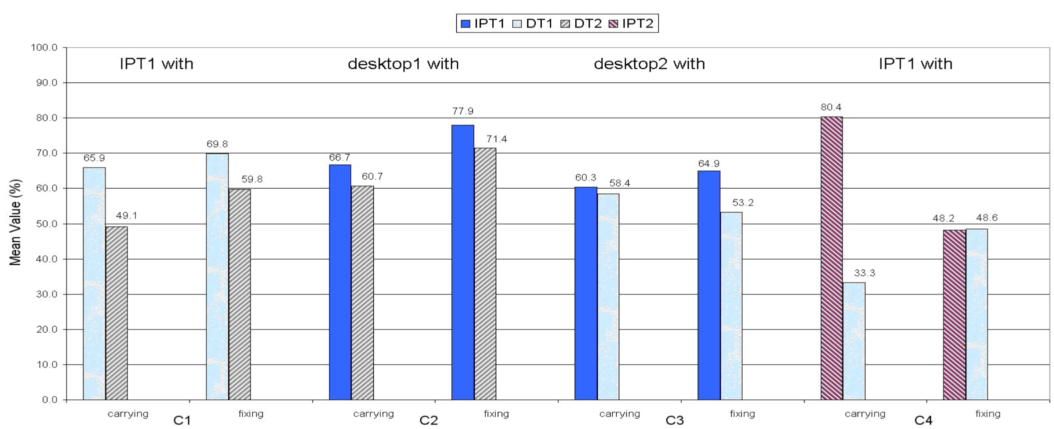


Figure 5-5: *Perceived collaboration while both carrying and fixing the beam*

5.6.6 *Presence, co-presence*

The participants were asked to answer three questions concerning the realism of sharing the task with another human being and of being in the same physical space.

"*Did the interaction appear realistic?*"

"*Did you feel that you were sharing the task with another human being?*"

"*Did you feel that both of you were in the same physical space?*"

The evaluation with ANOVA of the first two questions did not show any significant differences (p_{Q1}=0.621, p_{Q2}=0.699) between the conditions, although the high mean value (Table 5-8) indicates a perception of realism in collaboration. The third question showed a significant difference ($F(3,99)$= 2.82, MS_W=2.83, p=0.046) only in the IPT1-IPT2-DT1 trial. The higher mean value (M=72.9, SD=21.8) for the two IPTs compared to (M=54.8, SD=24.6) for the IPT2-DT1 connection demonstrate that immersive displays give a better feeling of sharing a working environment. Table 5-8 shows a summary of the questionnaire results for all three questions.

Table 5-8: *Questionnaire results relating to presence and co-presence*

Questions	**IPT1**		**DT1**		**DT2**		**IPT2**	
	mean	SD	mean	SD	mean	SD	mean	SD
interaction realism	51.6	24.6	49.5	21.5	53.6	24.5	61.4	16.6
co-presence	65.1	20.3	71.4	22.8	63.6	23.4	73.2	24.7
shared physical space (see Table 5-3 for user 2&3, depending on condition)								
with user 2	68.3	21.7	69.0	21.0	59.5	21.0	72.9	21.8
with user 3	64.3	26.6	70.2	21.5	54.8	21.0	54.8	24.6

Intensive involvement in a collaborative task can create a feeling of presence and co-presence. The results should confirm this yet fail to do so for asymmetric device combinations (non-immersive desktop vs. immersive CAVE-like). It is likely that this is due to the crude visual implementation of the Virtual Gazebo and the limited support of non-verbal cues. Nevertheless, some of the participants pointed out that: "[The environment provided a] fantastic experience of interacting with another person from a distance" and "I could really feel as if I was part of a team". In contrast, when two immersive users collaborate they have a high feeling of co-presence. This confirmed findings of earlier co-presence research [Axelsson et al., 1999; Schuemie et al., 2001].

5.6.7 Social feeling

The questions "*Did a lack of social feeling make the task harder?*" and "*To what extent did the interface hamper the task?*" were further asked. Looking at the first evaluation for a "lack of social feeling", an ANOVA revealed that there was a significant difference between the conditions ($F(3,50)$= 3.99, MS_W=2.44, p=0.013). A posthoc test showed that the difference was significant between the IPTs (IPT1: M=55.6, SD= 26.0 and IPT2: M=50.0, SD=22.6) and the desktops (DT1: M=75.8, SD=13.5 and DT2: M=72.6, SD=23.2), Figure 5-6. The figure demonstrates also the results about the interface and the significant difference between the conditions ($F(3,89)= 5.35$, $MS_W=2.48$, $p=0.002$). A posthoc test reveals that the difference lies between IPT1 (M=48.3, SD=23.6) and DT1 (M=72.0, SD=18.1).

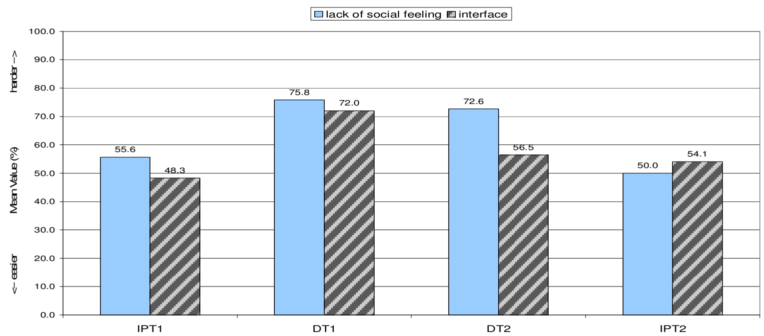


Figure 5-6: *Did one of the circumstances make the task harder?*

As the high values for DT1 and DT2 indicate, desktop subjects reported to have more difficulties to feel that they are in a social context. It can be argued that in such a highly collaborative task it is important that the participants feel as if they are working in a team and find themselves in a social context. When involvement in the task increases, it becomes essential to understand a partner's actions and to synchronise the joint interaction. Therefore, it is no surprise that the more engaged DT1 user (with audio to both collaborators) found the task harder than DT2 (with limited audio) due to difficulties with the different interfaces. An easy to use interface for this kind of interaction is difficult to design yet essential for an uninterrupted workflow. Some of the users pointed this out in the post-interview: "[it was] difficult working with/inside 3d environment in 2d" or complained of "the need to fiddle with the closeness of the controls".

5.6.8 Perception of verbal and non verbal influences

Further questions were asked concerning aspects of verbal and non-verbal contributions to the task. The participants gave their opinion on "*To what extent did each of the following contribute to task performance?*" in relation to the verbal communication, non-verbal communication, shared objects and the environment. The evaluation with ANOVA showed that there is a significant difference between the participants using the display type IPT1, DT1 and DT2, Table 5-9. However, there was no significant difference for C4, the IPT1-IPT2-DT1 trial.

Table 5-9: *ANOVA results for Verbal and non verbal contribution*

Display type	ANOVA results (α=0.05)	posthoc test (Tukey) shows differences between:
C1-IPT1	$F(2,48)= 3.78$, $MS_W=3.10$, $p=0.014$	verbal (M=70.7, SD=28.4) non verbal (M=45.4, SD=30.0)
C2-DT1	$F(2,34)= 5.06$, $MS_W=1.96$, $p=0.005$	verbal (M=76.2, SD=22.2) non verbal (M=45.7, SD=17.6) shared obj (M=52.9, SD=25.2)
C3-DT2	$F(2,30)= 3.46$, $MS_W=2.05$, $p=0.025$	verbal (M=61.9, SD=20.5) non verbal (M=37.7, SD=9.6)
C4-IPT2	$F(2,19)= 0.89$, $MS_W=4.16$, $p=0.457$	none

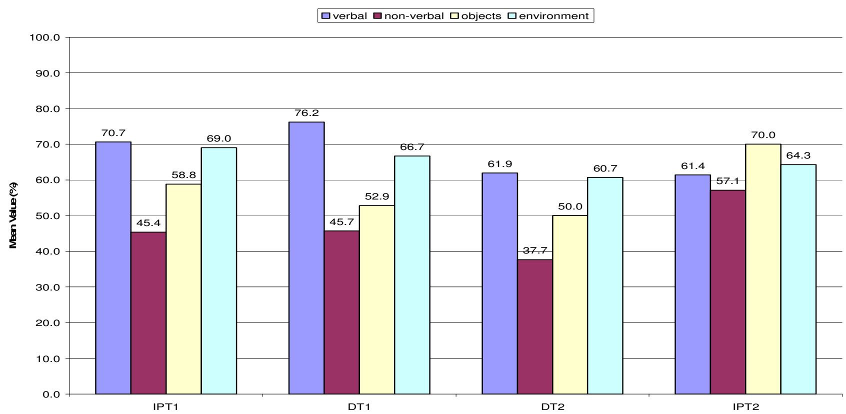


Figure 5-7: *To what extent did SHC contribute to task performance?*

The results of Table 5-9 show clearly that the participants of conditions one-to-three found verbal communication significantly more important than non-verbal communication. This was different for condition four where, as already mentioned, the two immersed (symmetric) users left the desktop user out of most of the activities. Under these circumstances, no significant difference between verbal and non-verbal communication could be found. In addition, shared objects were found to contribute more highly (M=70) for C4 than under the conditions C1-C3 with an average mean value of M=55. However, for all four conditions the interviewed user perceived the environment as similarly important (M=64).

It should be noted that all participants were observed to partly ignore the ability to talk to each other for synchronising their action, even though they knew each other. Although the observers tried to encourage them to use this facility, a lively discussion was seldom observed as one person pointed out: "[There was a] lack of verbal communication despite equipment provided". These behaviours were observed for the two different methods of communication provided for the participants as on the one hand a technology restricted microphone connection and on the other a non-restricted direct communication. It could be argued that this reluctance to verbal communication is counter productive to the highly collaborative task. There are various reasons to consider that could have caused this behaviour. One reason could be that people did not have support for direct eye contact or facial expressions which contribute to trustful communication. For example, in the real world people tend to look at each other to see whether their communication partners are listening, show interest and agreement/disagreement. Another possible explanation could be that users felt themselves to be observed and because of this didn't want to talk too much. Lastly, participants had to handle a task at the time, thus they were more concentrated on analysing the situation and acting accordingly as well as trying to use an unfamiliar interface.

Desktop users can manipulate distance objects through space (using ray-casting) and without changing their own position whereas the immersed user can only manipulate the object by moving to it and have it within his/her reach. This desktop behaviour is a common feature for CVEs (see Chapter 3) and designed to reduce the movements of the avatar, but also to simplify the object handling. However, as found in previous studies [Hindmarsh et al., 2001], this makes it more difficult for other users to see the relationship between the acting user and his/her object of interest. The same conclusion can be drawn for such a task as the Virtual Gazebo. People became confused about who was doing what. It could be observed that participants were surprised when another desktop user interacted with an object of their own interest. They started to ask questions such as "Are you taking the metal joiner? I thought I should take it".

5.6.9 Observation of user behaviour

During the user trials a number of observations were made in addition to recording user's opinion. They reveal that desktop users had considerable difficulty to learn the DIVE interface even though the most necessary commands where taught. In contrast immersive users grasped the interface relatively quickly and follow-up questions (regarding the interface) were not asked. Comments from desktop users like "Hard to manipulate objects, shift and middle click is hard to use", "Using a 2D device obviously hindered the process" or "interface controls not very intuitive" were quite common.

In addition, observations showed that desktop users worked from distant places relative to the workspace (e.g. 20m away from the construction site) to increase their field of view (FoV) and that subsequently they used left-right rotations to assess what the other participants were doing. In contrast immersive users had only to rotate their head. However, in the course of the trial immersive users tended to work on all walls, eventually facing the open end where they need to use joystick rotation to reorient themselves. A five-sided CAVE would avoid such problems and it would be interesting to see how this could influence an immersive user's performance as well as perception of presence. Nevertheless, immersive users still had very good 3D perception compared to desktop user which also was noted by participants "Only the person with the display that is tracked sees the world accurately".

Unfortunately, sometimes inconsistencies in the application could be observed where objects would not be securely fitted or created screws that could not be seen by all participants. This was due to overload of the network and a detailed analysis of these event losses as well as strategies to avoid or prevent such loss can be found in Wolff et al. [2004]. Subsequently this affected the closely-coupled collaboration between users "verbal communication and 3D object appearance is affected by network congestion" as verbal commands arrived earlier than visual representations of actions. On one occasion the audio failed during the trial and participants had to resort to their knowledge of the application goal and non-verbal communication, or as one user puts it, as a positive comment, "The loss of two-way communication forced us to work around this fact to achieve the goals".

Even though non-verbal presentation of an immersive user was limited to head, arm and body movements, it still had the desired effect on collaborators as reflected by one user "cave user's movements very easy to see". In contrast, immersive user complained that "the other users on consoles took a long time to do tasks" and it was "difficult to see what desktop users were doing". Obviously this was to be expected as non-verbal communication was very limited on desktops as the nature of their interface forced them to work from a remote distance (widening their FoV). Yet most users enjoyed the experience and stated that it was a "fantastic experience of interacting with another person from a distance" and "trying to work together from different situations is fun" or simply "The possibilities!!".

5.7 Summary

A degree of co-presence has long been supported by CVEs, however, the realism of shared object manipulation has, in the past, been hampered by interface and network delays. We have shown that a task requiring various forms of shared object manipulation is achievable with today's technology. This task has been undertaken successfully between remote sites on many occasions, sometimes linking up to three remote CAVE-like displays and multiple desktops. Detailed analysis has focussed on team performance and user evaluation.

5.7.1 Team Performance

Using the Virtual Gazebo application, novice users adapt quickly to remoteness of peers and the interface. Typically after three sessions their performance efficiency doubles, approaching that of expert users. Immersive users can undertake most parts of the task far more efficiently than their desktop counterparts. The Virtual Gazebo task requires collaboration at numerous points. This means that a faster user must often wait for the slower one to catch up before beginning the next step. Schroeder et al. [2001] found that the perception of collaboration is affected by asymmetry between users of the different systems. Our results show that the time taken to complete a collaborative task is also affected. When roles in the Virtual Gazebo task are ill-defined, the performance of the team approaches that of the weakest member (see 5.5.2). However, the performance is greatly increased when the immersed user undertakes the more difficult part of every task.

5.7.2 User evaluation

The user evaluation of the two distinct methods of object sharing is summarised in Table 5-10.

Table 5-10: *Summary of user evaluation of two distinct methods of object sharing*

	Same attribute	**Distinct attribute**
Contribution	IPT > Desktop	IPT = Desktop
Hindrance	IPT = Desktop	IPT = Desktop
Collaboration	IPT: IPT > Desktop Desktop: IPT = Desktop	IPT = Desktop

The findings of the questionnaire confirm that the perception of contribution is affected by asymmetry (immersive vs. non-immersive) of linked displays when *carrying* a beam. However, this is clearly *not* the case when *fixing* a beam (Figure 5-8). This suggests that the interface plays a *major* role during the sharing of an object's attribute and a *minor* role when sharing an object through distinct attributes. Surprisingly, neither the interface, nor the form of object sharing, is perceived to affect the level to which the remote user hindered the task. This appears to contradict the results of the performance analysis above. From the perspective of immersed users, collaboration is considerably easier with a symmetric user and the next Chapter 6 will look into this. However, a desktop user found the type of remote display to play little part in the level of collaboration. Chapter 7 discusses in more detail the influence of displays and interfaces on the performance and perception of a single user. It will demonstrate that there is a significant difference in the perception between low and high immersive displays which influences behaviour and interaction.

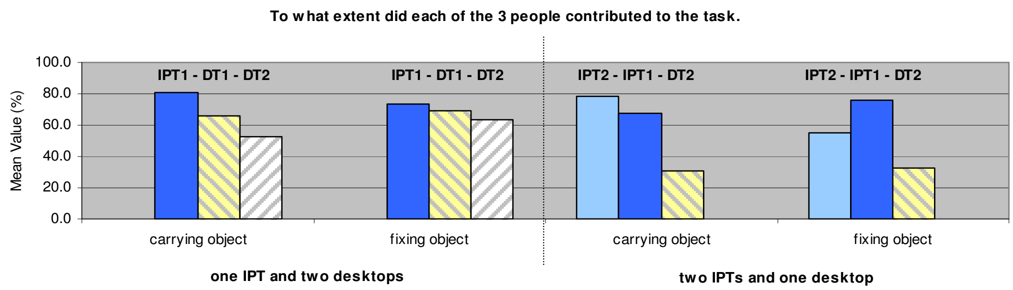


Figure 5-8: *Perceived performance results comparing IPT and desktop displays*

5.7.3 Workflow

The Virtual Gazebo benchmark is a highly collaborative application which involves all forms of closely-coupled collaboration and for a successful completion it is important to ensure a continuing workflow. It is not only important to enable users to manipulate the objects intuitively, but also to support them with different communication methods. The evaluation results also show that social communication is essential for such a task (e.g. to overcome issues with interface and application) and that each aspect of this, such as verbal, non-verbal communication or interaction with and around objects, should be involved. The results demonstrate that verbal communication is important for synchronising interaction and therefore completing the task successfully. Like in a related study by Hindmarsh et al. [2000] the observations and results show that desktop restrictions and remote behaviour do influence peoples workflow as it makes it more difficult for other users to see the relationship between the acting user and his/her object of interest. In contrast, the activity of immersed users is easier to interpret as they directly interact with objects or other people. However, the question that remains is how much non-verbal support, for example gestures, can and should be implemented without complicating the users interface.

Chapter 6

6 Testing the benchmark with two immersive displays

Continuing from Chapter 5, the benchmark application is used for another set of trials and its results are presented in this chapter. Having lessons learned from the previous trial, a slightly modified setup was used to allow for better observation and evaluation of user behaviour and performance. A detailed description is given in the following sections as well as a reflection on the motivation for the experiment.

6.1 Hypothesis-2: Immersive CVEs could support a seamless flow of collaboration and result in a performance efficiency increase

Previously in Chapter 5, user trials were mainly conducted between asymmetric (immersive vs. non-immersive) display setups. The chapter demonstrated that immersive users take a leadership role and that desktop users struggle with their interface during complex tasks such as moving and positioning an object in the 3D environment. This has an impact on the workflow and performance of a closely-coupled task. The results also showed that two immersive users outperformed a third connected desktop user. Therefore the hypothesis is that users connected with only immersive (symmetric) displays can work seamlessly and increase performance of a closely-coupled task. To test this hypothesis a new user trial was conducted which is now described, assessed and discussed.

The benchmark, the Virtual Gazebo, and its development is described in detail in the Chapter 4. Section 6.2-6.4 introduces the task and the setup for the various displays. The results are given in section 6.5 & 6.6, discussed thoroughly in relation to previous studies and finally summarised in section 6.8.

6.2 Task breakdown

Two immersive CAVE-like displays were connected from the University of Reading and the University of Salford. The experimental task introduced in chapter 5 showed that users start their structure at different places, which meant that detailed comparison of different trials is more difficult. Therefore, for the following trial the application was provided with a base-structure (two vertical beams) which participants had to complete (see Figure 6-1). The detailed breakdown of the new task is given in

Table 6-1, where an example is shown of how two users might construct this simple structure. The new task still required all forms of object sharing (table 4-1, pp. 62) along with varied human communication.

Table 6-1: *New detailed task breakdown showing example collaboration*

Sub-task	Description	User 1	User 2
ST1	Place T joiner	Fetch T joiner and hold it in place on first upright beam	
ST2	Drill hole	Hold the T joiner in place	Fetch drill and drill a hole through foot and T joiner
ST3	Insert screw	Hold the T joiner in place	Fetch a screw and insert it in the hole
ST4	Tighten screw	Hold the T joiner in place until screw tightened	Fetch screw driver and tighten screw
ST5	Place T joiner		Fetch T joiner and hold it in place on second upright beam
ST6	Drill hole	Fetch drill and drill a hole through foot and T joiner	Hold the T joiner in place
ST7	Insert screw	Fetch a screw and insert it in the hole	Hold the T joiner in place
ST8	Tighten screw	Fetch screw driver and tighten screw	Hold the T joiner in place until screw tightened
ST9	Carry beam	Fetch carry tools and use one to lift each end of the beam. When both ends are lifted, carry the beam to the T joiners	
ST10	Place beam	Place one end of the beam in a T joiner	And then lift the other end so that the beam fits in the other T joiner
ST11	Drill hole	Fetch the drill and drill a hole through a T joiner and beam	
ST12	Insert screw		Fetch a screw, insert it in hole
ST13	Tighten screw	Fetch a screw driver and tighten screw	

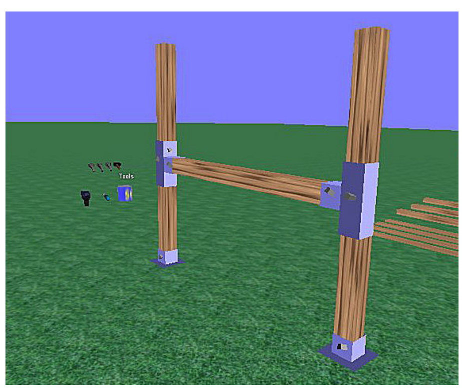


Figure 6-1: *The task target of a simple structure (vertical beams were already standing)*

The methodology for evaluating the task is explained both for team performance (section 6.5) and subject perception (section 6.6). Team performance measures the time taken to complete the task and each component sub-task. User evaluation details the responses to a questionnaire on the perception of collaboration. Observations and video transcripts made during the trials are discussed throughout this chapter.

6.3 Display configurations

The tests involved two immersive CAVE-like display configurations with the same field of view and similar user interfaces (see Table 6-2). The Salford IPT supported two handed tracking for additional non-verbal support and object manipulation was facilitated using a tracked joypad. To manipulate an object the user would press a button to grasp or select an object and only releasing the button would also release the object from a users grasp.

Table 6-2: *Display configurations*

Name	**IPT1**	**IPT2**
Location	Salford	Reading
Display	4 wall cubic display	4 wall cubic display
Input	Tracked Joypad with attached sensor	Tracked Wand incl. a Joystick
Tracking	Magnetic Flock of Birds	Ultrasonic/Acoustic Intersense IS900
Computer	4 pipes 12 dedicated processors SGI Origin 2000	2 pipes 6 dedicated processors SGI Origin 2000
Audio	Yes	Yes
Embodiment	Motion tracking	Motion tracking
Tracking filter	1cm / 200ms	1cm / 200ms
Observation	- video camera recording person in IPT - audio recording of verbal communication - observer recording behaviour and time - video-audio transcripts of three selected sessions	

Tests were undertaken over five days. During this period, typical network latencies from Reading to Salford were around 19ms. Similar to chapter 5 the DIVE CVE was used with activated event monitoring plugin and an event filter. An updated version of the Virtual Gazebo application was used were all object behaviours where no longer implemented as a Tcl/Tk script but as a C++ plugin. The advantage of this was a greater reliability of the application as the plugin execution was quicker and better optimised than the Tcl/Tk script.

Throughout the trials various measurements were taken. Tools for measurement were the above mentioned event monitor plugin (for path analysis) and time measurements of sub tasks. In addition, video and audio recording for both displays were taken as well as detailed observations during the trial from an observer.

6.4 Test Conditions and Questionnaire

The perceived effectiveness of collaboration involving shared objects and the perceived effect of display type were investigated, using a user evaluation questionnaire. Fifty four teams of volunteers were formed to conduct the tests and all participants signed a consent form (see Appendix A.2). For IPT2 each trial involved a new user while at IPT1 only ten users could be found and for all other sessions an expert user was used instead. Each user had at least 15min training before the trial to adapt to the user interface and benchmark application. Previous tests showed that after a short time of adaptation users would perform nearly as well as an expert user, allowing to concentrate on the task and less on the interface. Users at IPT2 were not aware whether a novice or expert user (a member of the research team) was collaborating with them. Within every task, each user interacted through a distinct display device and was questioned after the trial on his/her perception of the effectiveness of teamwork.

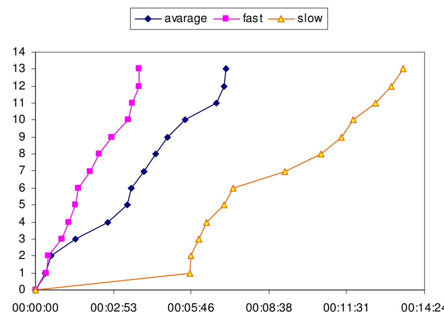
Similar to chapter 5 the questionnaire was aimed at ascertaining the user's subjective perception of collaboration, both generally and for each specific task. The questionnaire from Chapter 5 (Appendix A.2) was adapted and fine tuned to the new research aim. Answers could be given on a Likert-type scale [Sitzman, 2003] of 1-7, where 1 represented strong disagreement and 7 strong agreement. Those scales were later converted to percentage in order to allow for better and clearer comparison. Errors arising from a user's misinterpretation of a question were reduced by asking sets of related questions. For example, "to what extent did the two of you collaborate" was contrasted with "to what extent did each user hinder the task". During the analysis where differences were observed for those alternating questions, a further analysis is performed and documented below. Summary findings of the questionnaire (see Appendix A.1) are represented in the next sections of this chapter. A special accent is given to questions related to social human communication and the shared manipulation of objects.

6.5 Team Performance

Team performance was measured both in terms of time taken to complete the task and each component sub-task, in order to gauge the support for collaboration offered by immersive CAVE-like displays. The teams were left to determine their own organisation of roles in a natural way as the task progressed. The only constraint was the order of the sub-tasks ST1 to ST13 as described in Table 6-1. While evaluating the user trials it emerged that sessions could be categorised as fast, average and slow sessions. The overall mean time for all sessions was 6.6 minutes, yet a number of sessions were below (4-5min) or far above (8-15min) this average. An example of each session "type" is given in Table 6-3 with detailed timing for the task breakdown.

Table 6-3: *Timing for fast – average – slow sessions*

session	speed (% overall)	fast (33%)	average (37%)	slow (30 %)
	example	*46*	*20*	*43*
	Start	00:00:00	00:00:00	00:00:00
ST1	Place T joiner	00:00:23	00:00:21	00:05:43
ST2	Drill hole	00:00:29	00:00:34	00:05:45
ST3	Insert screw	00:00:59	00:01:29	00:06:02
ST4	Tighten screw	00:01:13	00:02:40	00:06:20
ST5	Place T joiner	00:01:28	00:03:25	00:07:00
ST6	Drill hole	00:01:36	00:03:32	00:07:18
ST7	Insert screw	00:02:02	00:04:01	00:09:14
ST8	Tighten screw	00:02:21	00:04:26	00:10:33
ST9	Carry beam	00:02:49	00:04:53	00:11:20
ST10	Place beam	00:03:26	00:05:32	00:11:45
ST11	Drill hole	00:03:34	00:06:41	00:12:36
ST12	Insert screw	00:03:50	00:06:59	00:13:11
ST13	Tighten screw	00:03:51	00:07:04	00:13:37
overall mean, SD in min		4.6 , 0.5	6.4 , 0.5	9.8 , 2.3



$F(2,54)= 64.61$, $MS_W=1.72$, $p=0.000$

An ANOVA over the data by session speed shows a significant difference between all three session types ($F(2,54)= 64.61$, $MS_W=1.72$, $p=0.000$). Further examination revealed that each "type" made around a third of all trials and that a main reason for the time differences lies in the degree of difficulties users experienced during a trial. For example, in slow sessions many users experienced difficulties to pickup or orient objects due to interface and/or application issues (see section 6.6.3). A visualisation of Table 6-3 shows that for a fast session interaction continues, while slower sessions contain sections of "interruptions" visible as longer durations between subtasks.

6.6 User Evaluation

For the analysis of the questionnaire the statistical approach of analysis of variance (ANOVA) was used to verify the significance of variations recorded, just in the same way as previously in Chapter 5.6. A posthoc Tukey test was applied if a significant difference could be found to clarify between which groups those differences appeared.

The evaluation is split into two categories, by display system and by session speed. The former may show significant differences in user perception of the interface and its contribution towards the task, while the latter may help to understand the influence of interruptions and problems during the task. Participants were allowed to complete the task in their own time and only during the user evaluation a pattern emerged which supported an evaluation by execution time (session speed).

Note: For presentation reasons the table attachment below is not further attached to any table in this chapter yet used throughout.

Where:	α is the limit of significant deviance MS_W is the mean square within groups F(a,b) is the variance between groups / MS_W p is the actual deviance, with four decimal places M is mean & SD is standard deviation	significant difference between: and as verified by the posthoc test (Tukey)

6.6.1 *Contribution during concurrent object sharing*

The first two questions of the questionnaire were "*To what extent did each person contribute to the task while fixing / carrying a beam?*". An analysis of variance (ANOVA) showed that there was no statistically significant difference between either question or categorisation, see Table 6-4 and Table 6-5.

Table 6-4: *ANOVA results for user contribution for oneself and the other user*

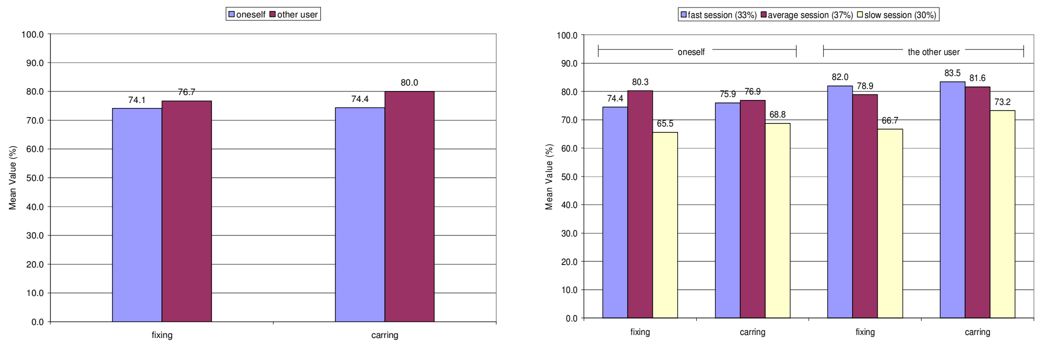
session speed	**oneself**		**other user**		**ANOVA results (α=0.05)**
	mean	SD	mean	SD	
… fixing …	74.1	19.4	76.7	19.6	$F(1,116)= 0.50$, $MS_W= 1.86$, $p= 0.479$
… carrying …	74.4	21.2	80.0	18.5	$F(1,116)= 2.48$, $MS_W= 1.97$, $p= 0.118$

Table 6-5: *ANOVA results for user contribution during fast – average – slow sessions*

session speed	**fast**		**average**		**slow**		**ANOVA results (α=0.05)**
	mean	SD	mean	SD	mean	SD	
oneself							
… fixing …	74.4	17.4	80.3	18.3	65.5	20.9	$F(2,54)= 2.93$, $MS_W= 1.71$, $p= 0.062$
… carrying …	75.9	17.1	76.9	21.9	68.8	24.6	$F(2,54)= 0.75$, $MS_W= 2.22$, $p= 0.475$
the other user							
… fixing …	82.0	16.1	78.9	19.5	66.7	21.4	$F(2,54)= 3.02$, $MS_W= 1.74$, $p= 0.058$
… carrying …	83.5	15.7	81.6	19.2	73.2	20.8	$F(2,54)= 1.49$, $MS_W= 1.67$, $p= 0.234$

The overall contribution throughout was perceived as good with high values between 65 to 83 percent. However, the results show a tendency of lower perceived user contribution during slow sessions compared to average or even fast sessions (see Figure

6-2b). It could be argued that this is due to difficulties during the trial. The more a user struggles with a problem the lower they judge their contribution and success.



a) *categorised by user* **b)** *categorised by session speed*

Figure 6-2: *Perceived user contribution during concurrent object sharing*

6.6.2 Performance and collaboration during concurrent object sharing

A second set of questions asked how users "performed" during the task and "collaborated" with each other. Firstly looking at the differences between displays shows that a significant difference for concurrent object manipulation of distinct attributes (fixing) can be found (Table 6-6). No such significance could be found for the concurrent object manipulation of the same attribute (carrying). However, the data show that for all cases, the IPT1 was perceived to collaborate and perform less than IPT2. The main differences between displays were the tracking system and interface. Although users from both sides felt that the interface hampered the task, the next section 6.6.3 shows no significant difference between sites. However, observations told a slightly different story. At the Reading IPT *two* cables from the tracking system and interface were hanging from overhead which allowed for relatively free movement within the CAVE-like display. In contrast, at the Salford IPT *four* cables were attached to the user from the ground and participants had to be careful not to stumble over them. The latter could even be observed by their remote partner at the Reading location "the other user stood on a cable, which interfered with the operation of fixing the t-junction in place". As unhindered movement within the display is more important during a fixing task it could explain the statistically significant difference between displays. In addition, it could be argued that because the Salford IPT had both hands tracked, the Reading participant (IPT2) had more non-verbal information from their partner available (see Figure 6-9), which IPT1 used among other things for guiding object positioning. The same was not possible for the IPT2 user, as only one hand was tracked. This could have led to a different perception of how the two partners collaborated and performed, as one partner seemed to get more help than the other. Again this would have been more important during fixing of objects as for moving (carry beam).

Table 6-6: *ANOVA results for user performance and collaboration for IPT1 and IPT2*

session speed	**IPT1**		**IPT2**		**ANOVA results (α=0.05)**
	mean	SD	mean	SD	
How well have you and the other person **performed** the task of ... an object together					
… fixing …	66.1	7.4	80.2	16.7	$F(1,56)= 4.55$, $MS_W= 1.32$, $p= 0.037$
… carrying …	76.2	12.4	81.6	16.0	$F(1,57)= 0.63$, $MS_W= 1.28$, $p= 0.429$
To what extent did the two of you **collaborate** while ... an object together					
… fixing …	71.4	18.7	86.6	11.0	$F(1,57)= 9.11$, $MS_W= 0.789$, $p= 0.004$
… carrying …	77.8	16.1	86.0	14.3	$F(1,58)= 1.74$, $MS_W= 1.17$, $p= 0.193$

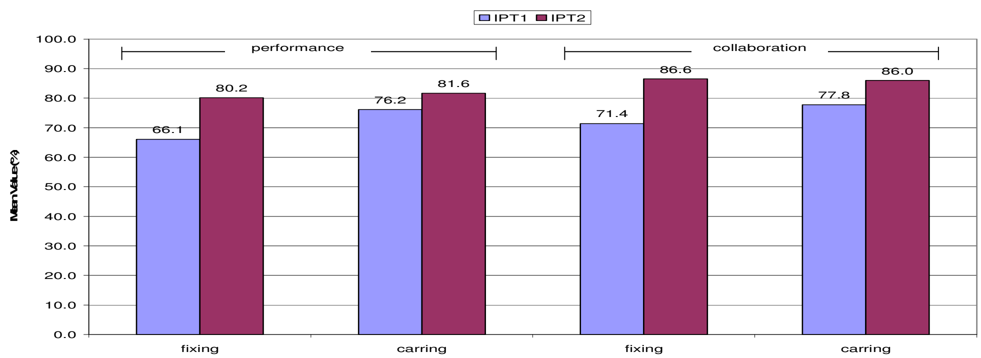


Figure 6-3: *Performance and collaboration categorised by IPT*

Further analysis of the data by session speed shows a significant difference between fast and slow sessions for the perception of performance (Table 6-7). No statistical difference could be found for the perception of collaboration. Collaboration is how people interact and work together, while performance represents how people contribute and successfully achieve a task. This is supported by the results, as people independent of their session speed nevertheless worked together and perceived such collaboration. On the other hand, interruptions of the workflow during slow sessions influence the perception of one's performance, which again is supported by the statistical results.

Table 6-7: *ANOVA results for performance and collaboration during fast – average – slow sessions*

session speed	**fast**		**average**		**slow**		**ANOVA results (α=0.05)**
	mean	SD	mean	SD	mean	SD	
How well have you and the other person **performed** the task of ... an object together							
… fixing …	86.5	13.4	77.6	15.4	71.4	17.9	$F(2,54)= 4.16$, $MS_W=1.16$, $p=0.021$
… carrying …	86.5	13.2	83.0	13.3	72.3	16.1	$F(2,54)= 4.37$, $MS_W=1.04$, $p=0.018$
To what extent did the two of you **collaborate** while ... an object together							
… fixing …	85.0	9.8	86.4	14.6	82.1	13.3	$F(2,54)= 0.50$, $MS_W=0.809$, $p=0.607$
… carrying …	85.7	14.3	86.4	13.9	83.2	16.2	$F(2,54)= 0.24$, $MS_W=1.06$, $p=0.789$

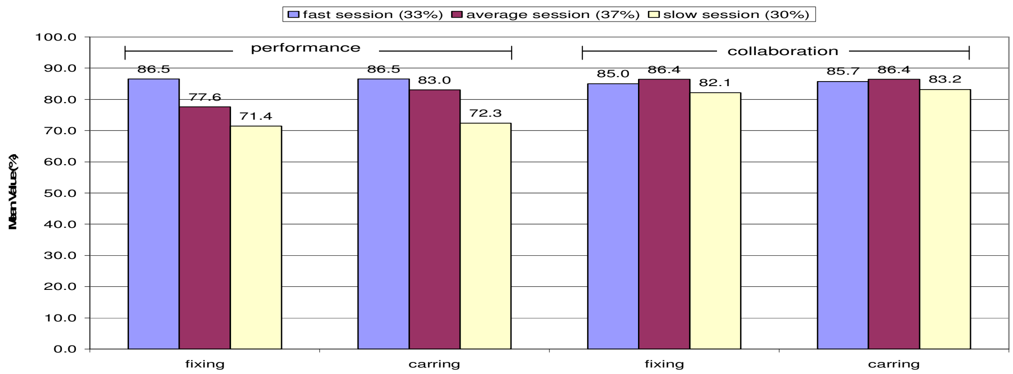


Figure 6-4: *Performance and collaboration categorised by session speed*

6.6.3 Hampering influences during closely-coupled collaboration

A further set of three questions tried to understand what hampered the interaction and collaboration during the trial. An ANOVA shows no significant difference between the displays (Table 6-8). System level issues like delays or inconsistencies were perceived to have no influence on the tasks. This was to be expected as no such problems occurred during the trials. However, the data demonstrate comparably high values for the hampering of the interface which is supported by observations where people had to adapt to move with a joystick or experienced problems in grasping objects. Here and in later discussions the interface seems to be a potential problem for closely-coupled collaboration. This already emerged in trials considering collaboration between desktops and CAVE-like IPTs (see Chapter 5.6.4), where desktops were significantly more difficult to handle than IPTs. This means that further consideration should be taken into the interface design and various research groups are investigating this issue. For example, users could use PinchGloves[2] to grasp objects without holding any device in their hand and use special designed Step-in-place platforms to provide free movements [Bouguila et al., 2002; *The VirtuSphere*, 2004]. Participants mainly complained about the cables, which seemed to have an influence on their perception of performance and collaboration (see above 6.6.2). A planned wireless interface and tracking system may resolve some of the issues, although an intuitive interface (no joystick) is still desirable.

[2] Pinch® Glove by Fakespace Systems

Table 6-8: *ANOVA results for hampering influences for IPT1 and IPT2*

session speed	**IPT1**		**IPT2**		**ANOVA results (α=0.05)**
	mean	SD	mean	SD	
To what extent did **the interface hamper** the task of … an object together					
… fixing …	50.0	20.2	43.4	22.0	$F(1,56)= 0.52$, $MS_W= 2.35$, $p= 0.475$
… carrying …	42.9	20.2	43.5	21.2	$F(1,56)= 0.01$, $MS_W= 2.16$, $p= 0.909$
How much have network induced **delays hamper** the task of … an object together					
… fixing …	17.9	6.6	22.7	14.3	$F(1,56)= 0.85$, $MS_W= 0.887$, $p= 0.361$
… carrying …	17.5	6.3	23.0	13.6	$F(1,57)= 1.36$, $MS_W= 0.799$, $p= 0.248$
How much have network induced **inconsistencies hamper** the task of … an object together					
… fixing …	20.4	7.6	31.0	23.1	$F(1,53)= 1.35$, $MS_W= 2.33$, $p= 0.251$
… carrying …	19.6	7.4	24.9	17.0	$F(1,54)= 0.69$, $MS_W= 1.25$, $p= 0.410$

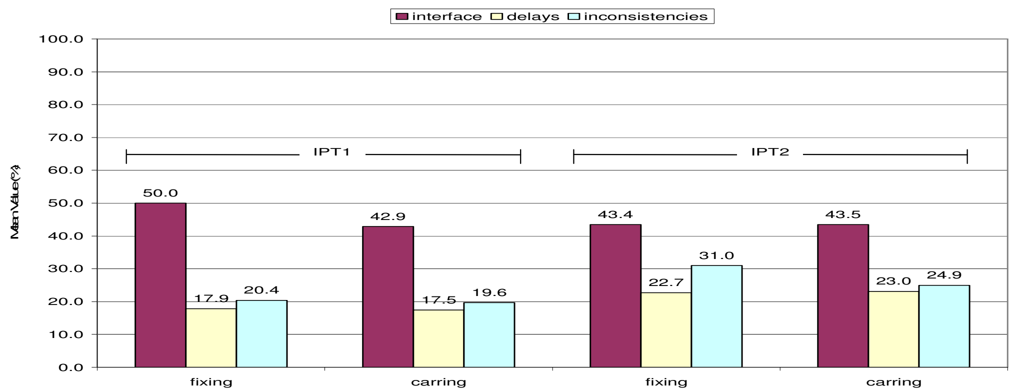


Figure 6-5: *Hampering influences categorised by IPT*

Looking closer to categorisation by session speed, the statistical analysis shows again no significant difference (Table 6-9). However, the data demonstrate lower perception values for fast sessions compared to slow session. This is to be expected as the workflow during fast session was seldom interrupted by issues caused by the interface. In the opinion of the users, delays did not hamper collaboration in any way which is consistent with the records where no network delays could be observed.

Note that inconsistencies have higher values for slow and average sessions compared to fast sessions. As mentioned before no system inconsistencies could be observed or recorded during the trials. It could be argued that the perceived difference is due to a different understanding of the question, whereas inconsistencies included problems like a screw does not stick (because a user forgot to drill a hole first) or difficulties to pickup an object.

Table 6-9: *ANOVA results for hampering influences during fast – average – slow sessions*

session speed	fast		average		slow		ANOVA results (α=0.05)
	mean	SD	mean	SD	mean	SD	
To what extent did **the interface hamper** the task of … an object together							
… fixing …	36.8	17.1	46.3	25.1	49.5	21.5	$F(2,54)= 1.66, MS_W=2.27, p=0.200$
… carrying …	41.3	18.0	39.5	19.6	50.0	22.1	$F(2,54)= 1.28, MS_W=2.11, p=0.287$
How much have network induced **delays hamper** the task of … an object together							
… fixing …	18.8	6.9	25.9	18.4	21.9	11.9	$F(2,54)= 1.35, MS_W=0.905, p=0.267$
… carrying …	18.8	8.5	24.5	14.4	24.1	15.4	$F(2,54)= 1.14, MS_W=0.827, p=0.327$
How much have network induced **inconsistencies hamper** the task of … an object together							
… fixing …	21.1	8.8	35.0	27.6	34.7	23.6	$F(2,54)= 2.52, MS_W=2.26, p=0.090$
… carrying …	20.3	14.1	23.6	14.1	30.5	20.1	$F(2,54)= 1.76, MS_W=1.23, p=0.182$

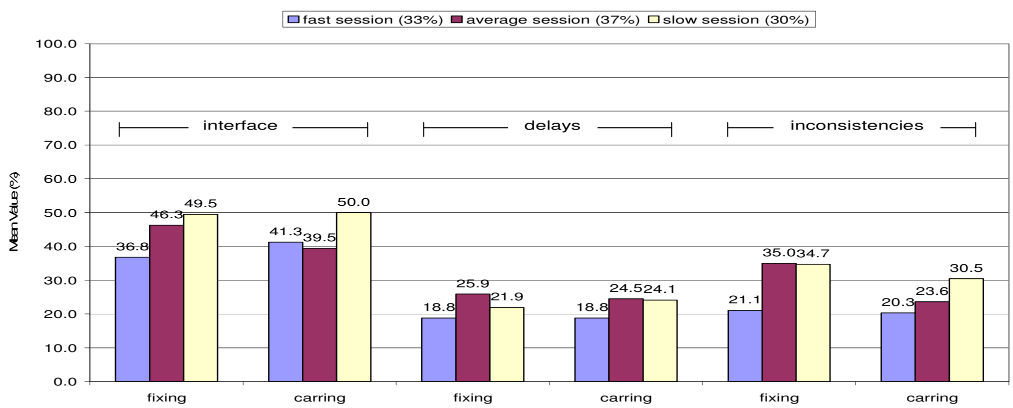


Figure 6-6: *Hampering influences categorised by session speed*

6.6.4 Perception of social interaction, presence and co-presence

More than seven questions were asked relating to the perception of being in the virtual environment and to share this space with another human. Figure 6-7 shows a summary of these questions and their mean values, all questions and their data can be found in Appendix A.2. No significant differences between the immersive displays could be found, but the data show high values for the perception of presence and co-presence. The graph also shows slight differences for some questions, none of which had statistical significance (Table 6-10). However, to the question "*Could you sense the emotions of the other persons?*" slightly higher values can be seen for IPT1 compared to IPT2. This is contradictory to the expectations, as IPT1 had both hand tracked and was better able to use non-verbal communication it would have been expected that IPT2 can better sense these cues.

Table 6-10: *ANOVA results for social influences for IPT1 and IPT2*

session speed	IPT1		IPT2		ANOVA results (α=0.05)
	mean	SD	mean	SD	
social interaction	74.9	19.1	69.7	21.3	$F(1,58)= 0.53$, $MS_W= 1.99$, $p= 0.472$
presence	77.3	18.9	68.1	21.9	$F(1,58)= 1.73$, $MS_W= 1.99$, $p= 0.194$
sense emotions	55.6	18.1	40.6	22.8	$F(1,58)= 3.16$, $MS_W= 2.45$, $p= 0.081$
realistic appearance	63.5	25.9	56.3	22.1	$F(1,58)= 0.79$, $MS_W= 2.47$, $p= 0.379$

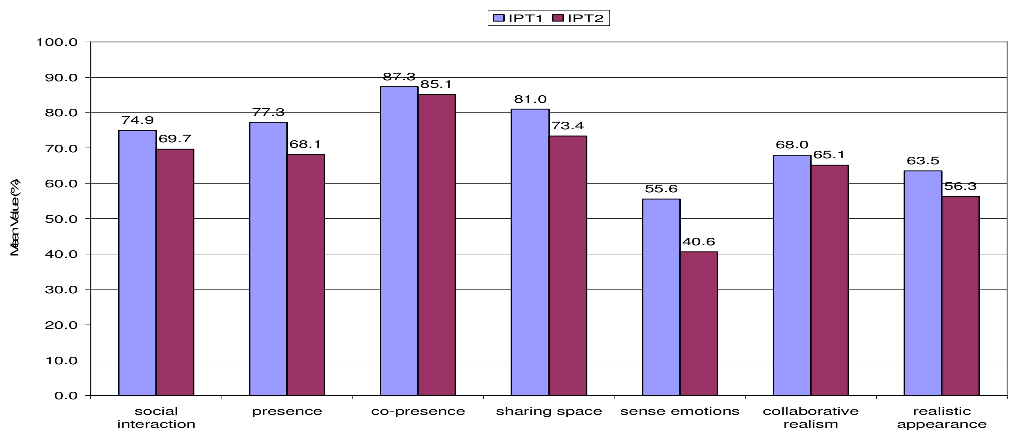


Figure 6-7: *Perception of social influences, presence and co-presence categorised by IPT*

Looking at the same set of questions categorised by session speed a different picture is emerging. Participants had independently of the progress a high perception of presence but most notably of co-presence. This demonstrates that the benchmark application and setup was successful in placing participants in a collaborative environment. Also a question regarding the level of social interaction showed high values for an average session. Although not statistically significant, the result could be interpreted as a response to more intense observed social talk. For example, participants during fast sessions did not have to talk too long as the progress was swift, on the contrary in slow sessions participants where busy trying to solve their problems with the interface or application. Average sessions on the other hand were a mixture (of successes and failures) and conversations between both participants were slightly more intense (see below extract of Transcript B-2).

Time	
52:35	Sara can see the screwbox near the other beam “the whole box is on the other side” “can you see it, yea”
52:43	Jim “yea, yea got it” he creates a screw and picks it up
52:47	Jim moves with screw to Sara and tries to insert it, but drops screw “up, shouldn’t have to do that” and <looks down>
52:50	Jim <bends> down and <picks up> the screw,
52:54	Sara “perhaps is better to take the other one” <she glances at him>
52:57	Jim: “what?, in the there” <insert successful> “ah yea, here we go”

Table 6-11: *ANOVA results for social influences during fast – average – slow sessions*

session speed	fast		average		slow		ANOVA results (α=0.05)
	mean	SD	mean	SD	mean	SD	
social interaction	66.5	21.1	75.9	20.3	64.5	18.5	$F(2,54)= 1.89$, $MS_W=1.82$, $p=0.162$
sense emotions	32.3	20.2	41.5	20.7	52.9	23.0	$F(2,54)= 4.24$, $MS_W=2.20$, $p=0.019$
realistic appearance	61.7	20.7	58.5	23.0	49.6	22.7	$F(2,54)= 1.44$, $MS_W=1.44$, $p=0.246$

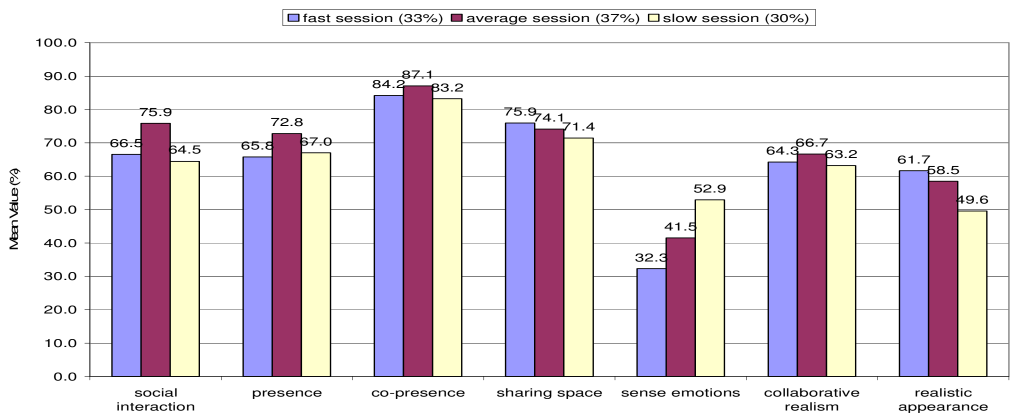


Figure 6-8: *Perception of social influences, presence and co-presence categorised by session speed*

Emotions are an important social resource in managing group collaborations. It is not only important to consider how technology may capture and communicate emotions but how it encourages them. Emotions are communicated through voice and body language, both of which can be captured to a limited degree and communicated within a CVE. Immersive CVEs track chosen aspects of body movement and use them to drive an avatar, which increases awareness for other participants. A feeling of presence contributes in allowing vivid experiences which can generate powerful emotions. In addition, the way users "get along" with the interface is influencing speed and performance of their collaboration. Combining this can have further implications on the application realism and transmission of user's emotions. Figure 6-8 shows lower user opinion for application realism when they encountered interface and application problems. At the same time users of slow sessions perceived a significantly higher emotional response than fast session users. It could be argued that this is due to higher verbal transmission of frustration (negative emotions) when problems persisted. This is supported by some evidence that negative moods and emotions are more easily transferred than the positive moods [Joiner, 1994; Tickle-Degnen & Puccinelli, 1999]. An example of this emotional response can be seen below in an extract of the Transcript B-3 (slow session). Here both users have tried for more than 3 minutes to fix a joint and after a few problems a build-up of negative emotions can be observed. Due to the previous problems they become uncertain about the action and start questioning how to proceed and when they believe they were successful another failure occurs.

Time	
44:06	(In the confusion both drop their item)
44:07	Shawn: “dropped it”
44:10	John: “oh no that’s ..”
44:12	While John is picking up the joint, Shawn moves back to the screwbox and gets another screw “I’ll get another one”
44:22	John is holding up the joint and asks “do you wanna really now” Shawn looks at him (with a screw in his hand)
44:28	Shawn: “I can see it now” sees the screwdriver in front of him and picks it up, he tries to fix the structure “yep” <pause> “I think you can release it now”
44:40	John lets go but it is not fixed and the joint drops down
44:41	Shawn: “NO”

6.6.5 Perception of verbal and non verbal influences

One question was concerned with aspects of verbal and non-verbal contributions to the task. The participants gave us their opinion on “*To what extent did each of the following contribute to task performance?*” in relation to verbal communication, non-verbal communication, shared objects and the environment. No significant difference between the displays could be found. However, the graphs show clearly the importance of verbal communication (Figure 6-9) as well as objects and the environment. Low values were given for non-verbal communication as the support by both immersive displays was limited, although higher values were given from IPT2 users. This is expected as users of IPT1 had both hands tracked and were thereby better able to show their intensions, give directions or use the second hand otherwise to the benefit of their collaborator at IPT2.

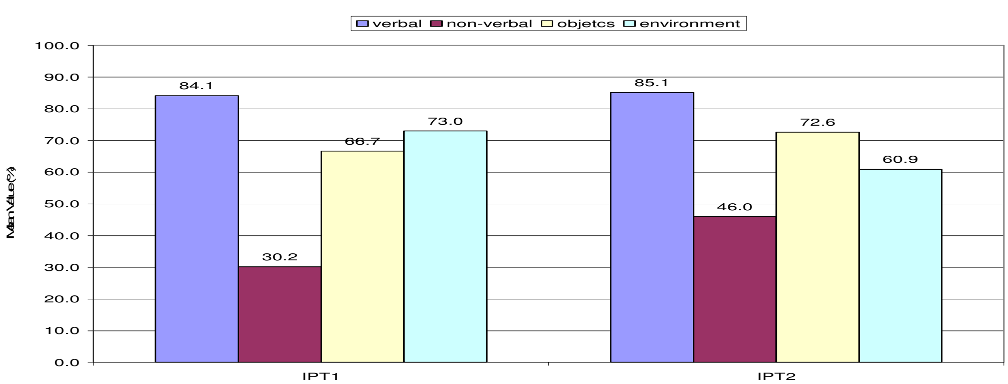


Figure 6-9: *Perception of SHC categorised by IPT*

Similar results can be seen between different sessions categorised by speed (Figure 6-10). Verbal communication has marginally higher values for average and slow sessions where, it could be argued, communication was used more often to transmit emotions. During fast sessions, however, interruptions of the workflow were limited and participants concentrated on finishing the task. As for all sessions talk was mostly limited to task relevant issues and it could be argued that because during fast sessions the flow between talk, interaction and navigation was "well balanced", these users laid slightly higher importance on non-verbal communication and objects.

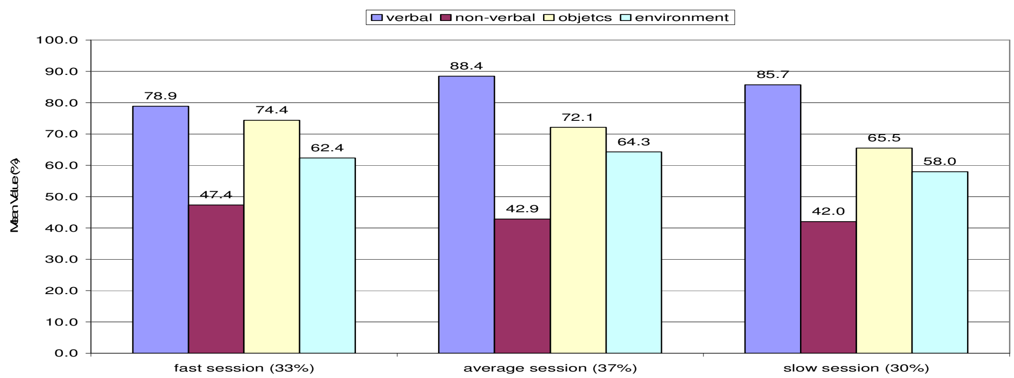


Figure 6-10: *Perception of SHC categorised by session speed*

6.6.6 Observation and Conversation Analysis for fast – average – slow sessions

All trial sessions were accompanied by observations and audio-video recordings. The video evidence was used to produce transcripts of selected sessions, one per category of fast-average-slow session. In contrast to section 6.5 about Team Performance, the timing for the transcripts are not normalised (starting with 00:00) to allow better comparison of the provided video evidence (see Appendix B).

A hypothesis to conduct the user trials discussed in this chapter was that immersive CVEs could support a seamless flow of collaboration [Roberts et al., 2006]. There is significant evidence to back up this hypothesis, see Transcript B-1 to Transcript B-3, of which a sample is presented below. However, there is also significant evidence against it. The reason for this is that some aspects of collaboration typically appear seamless, whereas other typically seems severally hampered by the technology or the user's ability to use the technology. For example, two participants are connected via the two IPTs, "Sara" has both hands tracked while "Jim" only has his right hand tracked. Both are novice users and had 15-30 min training to become familiar with the interface and application. Their goal of building the simple structure has progressed to the point that one t-joint was just fixed and the seconded t-joint now needs fixing. A convention is used to describe different activities of planning (P), agreeing (A) and doing (D). Planning can be described as “having an idea or plan of how to progress”, while agreeing is “mutual acknowledgment of each others action” and doing is a convention of “active behaviour like picking something up”.

Extract from Transcript B-2 showing fluent transition between planning (P), agreeing (A) and doing (D)

P	A	D	Time	Sara (non-native English speaker) in Salford has both hands tracked while Jim (native English) in Reading has only one hand tracked. Both are novice users and had 15-30 min training.
				Both participants just fixed their first t-joint and now have to fix the second one.
■	■		51:43	Jim: “are you going to hold up the other one this time?”
■	■		51:45	Sara: “ok”
■		■	51:49	Sara moves to the joint stack and picks up another t-joint
		■	51:52	Jim looks for the drill and picks it up, then moves with joystick back to have a look at both beams
■			51:55	Sara is on Jims front screen, Jim <glances> at her
■			51:58	Sara: rotates with joystick so that the constructions site is on her left CAVE wall
■	■		51:59	Sara: is moving towards the second beam, Jim who is still looking at her points with his hand (which holds the drill) to the beam “I guess it come right into the middle”
■		■	52:05	Sara <overshoots> the target, ends up at the first beam, is confused and looks around
■	■	■	52:09	In the mean time, Jim sees Sara near the beam and says “right get them about the same height” and <points with the dill> “so there”
■			52:12	Sara: “aeh where is it” (beam is on the open side of the CAVE), rotates and finds it
		■	52:21	Sara is moving in the CAVE to align the t-joint with the beam and Jim asks instructor “and then what do I press to use it again?” he just need to intersect the drill with the joint
		■	52:28	and now Jim is <doing> it “got it right”
	■		52:32	Jim: moving backwards “getting some; will see”
■	■		52:35	Sara can see the screwbox near the other beam “the whole box is on the other side” “can you see it, yea”
■	■		52:43	Jim “yea, yea got it” he creates a screw and picks it up
		■	52:47	Jim moves with a screw to Sara and tries to insert it, but drops the screw “up, shouldn’t have to do that” and <looks down>
		■	52:50	Jim <bends> down and <picks up> the screw,
■			52:54	Sara “perhaps is better to take the other one” she glances at him
■	■		52:57	Jim: “what?, in there” insert successful “ah yea, here we go”
		■	53:03	Jim moves to find the “ah screwdriver” and Sara follows his movements (looking with her head)
		■	53:08	Jim <overshoots> the tool “upp” and physically rotates to get into a better pickup position
		■	53:14	Jim picks up the screwdriver and moves back to Sara and the joint
	■	■	53:21	Sara looks at him “na”
	■	■	53:22	Jim fixes the structure “here we go”

All transcripts (see Appendix B) as well as this small section show the transition between different phases of interaction is fluent as is the seamless flow of conversation of both verbal and non-verbal aspects. This seamless flow of collaboration between distributed CVEs is hard to distinguish from the real world. However, this completely changes when inexperienced users attempt to pick up and carry objects. Some inexperienced users spend significant proportions of time "fumbling" with objects because they had difficulties using the interface. Users needed to point the virtual hand into the object and press a button on the wand/joypad which could be difficult if the virtual hand was not positioned right or the wrong button was used.

Another observation was that in the course of the trial immersive users tended to work on all walls, eventually facing the open-end where they need to use joystick rotation to reorient themselves (see above transcript at 51:58). A user commented that he "felt need for the 4th wall behind me in the CAVE few times while searching for objects". Furthermore, users complained about wires from the tracking and interface to be a problem as it hindered them to freely move and interact, "the cutting out of my glasses and the cables restricted or confused my movement". A five-sided CAVE with a wireless interface would avoid such problems and it would be interesting to see how this could influence an immersive user's performance.

The next subsections will describe various collaboration scenarios (see Chapter 4.4) and will take a closer look into observed user behaviour that was made during the user trials.

6.6.6.1 Scenario 1: Planning and Instruction

Planning is necessary to determine method and responsibilities. Instruction occurs when a person demonstrates to the other how to undertake a given operation, such as using a tool to fix two construction objects together.

The process of planning and instructing requires that everybody involved sees and hears the discussion. Verbal communication is essential to describe the upcoming task and to agree on locations and coming steps. Other cues, such as gestures, are widely used to point out directions and to underline the verbal communication. Earlier studies showed that simple embodiments contribute to the interaction and a more realistic humanoid avatar representation may support better collaboration [Bowers et al., 1996]. Thereby the faithfulness of the avatar gestures is as important as the realism of the environment.

Observations show that in the planning phase, users mainly stay close together or use body-centric gestures such as facing each other, while they discuss their next action. The use of CAVE-like displays supports this kind of user behaviour by allowing the user to turn and move naturally within the spatial context. Planning can also involve the use of objects. For example, when explaining how or where to use a tool, it is easier to take the object and demonstrate it. Using objects to describe an action is typical in real world interaction, even if the object is only used to mimic the action. When the environment is designed to support such communication, it contributes to the planning and also later to the task. For example, the use of different textures on similar objects can help to make verbal references to those objects. Apart from the environment, the technology has to be suitable as well and fully immersive displays would be the best choice as they allow orientating on and navigating a full 360º. In four-sided CAVEs, such as the one used in this trial, the open wall can mean that remote participants refer to something which is not immediately visible for the other user, resulting in additional orientation through rotation. This could occasionally be observed during the trial, however, in comparison to desktop users (Chapter 5, [Hindmarsh et al., 2000]) its frequency and impact was far less.

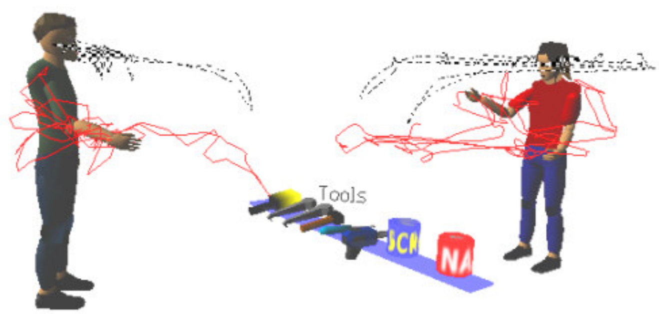


Figure 6-11: *Traced head (dotted line) and hand (solid line) events in a planning scenario*

From video footage taken during user trials, numerous nuances have been observed that link verbal and non-verbal communication while referring to objects and places within the environment (e.g. Transcript B-1, 42:55). For example, a user might point to an object and say "lets pick that up" and then turn and point to a place in the environment saying "and take it over there", or the user simply takes an object and tells the other user to do the same. Much of the non-verbal communication identifiable during planning consisted of turning, pointing and nodding.

Figure 6-11 shows avatar trajectories over 54 seconds from a selected dataset that visibly resembles the distinct roles between the collaborators. It illustrates how the two subjects moved their heads and hands, while discussing and planning the usage of the construction tools that the benchmark application provided. In this example, the avatar shown on the left side was explaining (instructor) and the one on the right was testing a tool following the instructions (listener).

6.6.6.2 Scenario 2: Working Separately

The team temporarily split to undertake separate but related tasks, such as gathering material objects, while keeping an eye on each other in case help is needed.

Most collaborative tasks involve some independent work. At such times, communication reduces naturally due to independence and concentration on one's own task. It is interesting to note that, at such times, small talk can replace work related verbal communication, and this seems to maintain the social feeling and increases the feeling of co-presence. Changes in level or type of activity or level of communication of the other person can signal a need to bring collaboration closer together again and communicate directly. It was often observed that people were looking over to see how their partner was getting on (e.g. Transcript B-1, 44:28 & 45:40) and offering assistance when necessary, for example, by fetching a tool for the other to use. Representing interaction with an object through natural body movements, driven by motion tracking, made such changes in activity easier to spot. Furthermore, the naturalness of view change offered by motion tracking as well as the wide field of view in a CAVE-like display, simplifies keeping a watch on others.

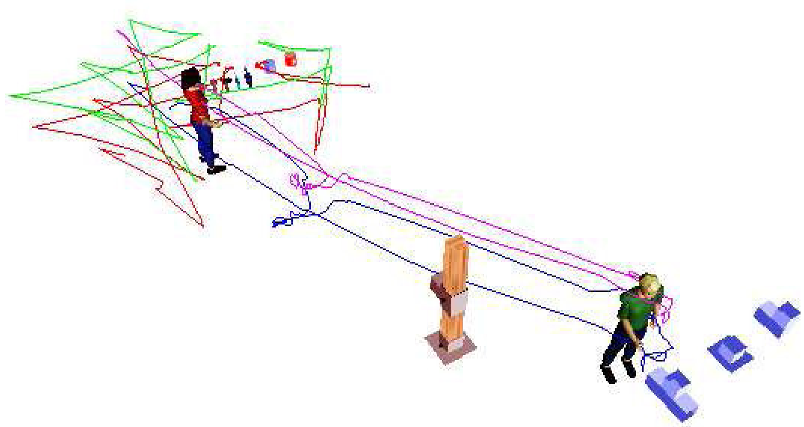


Figure 6-12: *Traced head and hand movements while working separately*

Observations during the trials have shown that some participants have used the joystick interface more frequently, while others made more use of the space within the IPT by physically walking towards close objects (e.g. Transcript B-2, 55:22). Figure 6-12 illustrates such an example of distinct user behaviour.

The ease and naturalness of collaboration between the test subjects was very noticeably improved above that observed in earlier trials that linked a desktop and CAVE-like display (Chapter 5). This was particularly apparent for parts of the task where both participants concurrently positioned one or more objects. Unlike previous desktop trials [Hindmarsh et al., 2000], delays were not often observed in identifying objects referenced through verbal and non-verbal communication. It is likely that the combination of wide peripheral vision and ability to glance using unconstrained head movement overcomes this problem.

6.6.6.3 Scenario 3: Moving an Object Together

For two remote users concurrently manipulating an object together is very demanding on the application and CVE [Wolff, 2006] as all participants ought to have a consistent and responsive view of the virtual environment. Here two users work together to carry an object to a given place in the environment (Figure 6-13). Simulated gravity ensures that larger construction objects cannot be lifted alone and simulated friction restricts dragging.

Supporting various forms of SHC is particularly important when people come to collaboratively manipulate an object. For example, when moving an object from one point to another, users must agree on roles and responsibilities as well as actions. Initially they must agree on where are they going, how are they going to get there, who will take the lead and what problems may be encountered. It is very important that the users can see which end of the object is picked up. Highlighting the manipulated part of the object supports this. Once underway, an agreement on speed and adaptations to path must be communicated and acted upon. In contrast to other distributed technologies (Chapter 3), as this chapter demonstrates, CAVE-like displays are well suited to support SHC in tasks such as collaborative manipulation of objects. This is because, in addition to supporting gesturing and voice as described above, each user interacts with the object in a spatially natural way with important aspects of body movement represented remotely with respect to the object.

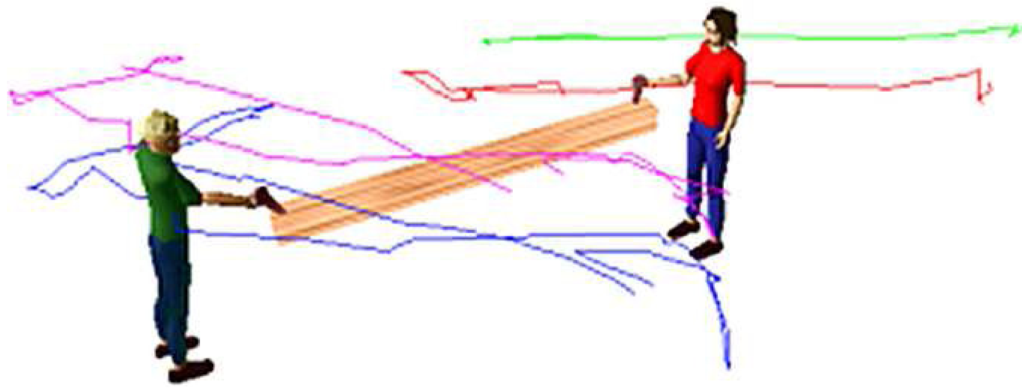


Figure 6-13: *Traced head and hand events while concurrently moving an object together*

During observations in the trials, the participants were again seen to adopt *leader* (avatar on the left side in front) and *helper* (avatar on the right side in the back) roles (e.g. Transcript B-1, 45:15). The leader is usually the one taking the initiative for further action. Here in Figure 6-13, the leading person usually directed the movement during carrying and selected the area to drop the object, whereas the helper was following and supporting the carry task.

Video footage shows that users combine speech with a complex set of gestures to communicate initial planning. This set includes nodding and shaking of the head and pointing, not only with the hand but also the head and body. For example a user may face his partner, point with the hand to the object and move the head between the two. Once the object has been lifted the users keep each other in sight, either by walking side by side or one in front walking backwards (Figure 6-13). As the primary hand is used to carry the object, gesturing plays less of a role and users make more use of body and head orientation (e.g. Transcript B-2, 53:40) as well as the object, the environment and verbal communication, in order to determine changes in movement and role.

6.6.6.4 Scenario 4: Assembling Objects Cooperatively

In the Virtual Gazebo application, objects are connected by holding them together, while drilling a hole through both and inserting and tightening a bolt. Typically, a construction will be built from the ground up, fixing one item in place at a time. Simulated gravity requires that one user must hold an object in place while the other fixes it (see Figure 6-14 for an example). The nature of the task requires that once again two distinct roles have to be adopted: a *leader* who directs the manipulation, and a *helper* who has a more supportive role. The avatar on the left shows how the helper has moved while holding the T-joiner for the leader on the right to drill a hole and finally fix the joiner to the beam.

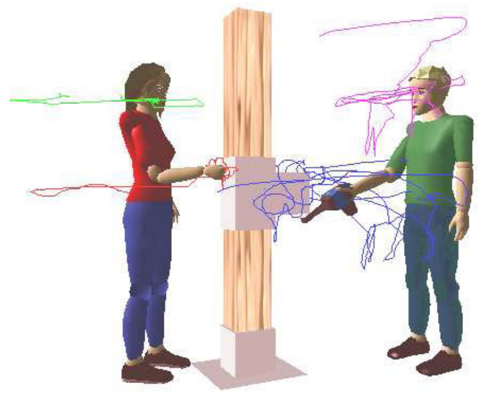


Figure 6-14: *Traced head and hand events while assembling objects*

Similar to scenario 3 of moving an object together, during an assembling task it is necessary to communicate with the other partner. Verbal communication helps to agree on the next step, but also provides the other user with further information. For example, a user positions an object while the other is told to gather the necessary tools to fix it. The first user can be of assistance by verbally directing how to fix it or informing the other person where a tool can be found ("look next to you"). Nevertheless, they can help each other by pointing out where to place an object (e.g. Transcript B-1, 42:55), how to use the object itself, and what to do next. Again users can support each other's work with a complex set of gestures. For example, nodding with the head or looking to the object instead of to the partner shows the other users agreement or a direction. As before, the technology of the CAVE-like display can support this kind of interaction. It allows the users to move with their body around the object of interests, without time consuming locomotion to get the right view position as would be required on a desktop. Furthermore, it is helpful when objects do support the communication between two users interacting around it, e.g. through a change of texture or other feedback when a "sub-goal" is achieved. Thereby, all participants can see how far the progress is and they may or may not continue with their work.

6.7 Discussion

For the benefit of better user evaluation the task had been modified for the CAVE-CAVE trial compared to the trial discussed in chapter 5. In addition, all users went through a 15min practice session to familiarise with interface and application. This makes it difficult to directly compare the time performance of both setups. However, if users were not hampered by their understanding of the application or usage of interface, the performance was comparable to expert users. If interaction was hampered it came mainly from difficulties while grasping objects, which in turn had different causes such as losing 3D vision when shutter-glasses lost the infrared signal, using the wrong button on the joypad or misjudging the distance to an object. Other problems can arise from the tracking system where cables could limit free movements within the CAVE-like display. Such difficulties with the interface can also have significantly negative impact on user's perception of performance and collaboration. This was in particular a problem for users of the Salford IPT where a "richer" yet "more constraining" tracking system (due to more cables) had a greater negative impact than for their Reading collaborators.

Compared to the trials of Chapter 5 observations could be made that showed a seamless flow of interaction and communication. This increased the perception of collaborative performance as questionnaire results show equally high or even higher values as Chapter 5 results. Yet performance differences were observed and categorised into fast – average – slow sessions. The main reason for slower sessions were due to problems of understanding the task sequence or difficulties with the interface in grasping objects or navigating. This resulted in giving significant different opinions during the questionnaire for contribution, performance and collaboration. Nevertheless, all participants had a high perception of social activity and presence in the environment as well as with others. In addition, transcripts show that at the beginning and end people focus on social interaction, while interaction via technical interfaces and the virtual representations play a more important role for task focused collaboration.

Interestingly, a significant difference could be found in the perception of emotions where negative emotions from slow session users were better perceived than positive emotions from users of faster sessions. In addition, as important as verbal communication is for closely-coupled collaboration the use of non-verbal communication should not be underestimated. Although non-verbal support by the IPT was limited and participants did not judge its contribution very high, observations show that gestures such as pointing, gazing, etc. were frequently used to support verbal cues like agreeing or moaning. Furthermore, observations show that participants respect each others personal space (see Chapter 2) by avoiding avatar collisions.

6.8 Summary

The previous chapter 5 concluded that immersive users take a leading role in closely-coupled collaboration and that they outperform desktop users within the same virtual environment. The workflow was interrupted by application and interface problems as well as orientation issues of desktop users. This raised the question of what would happen in a pure immersive setup. The hypothesis was that immersive CVEs could support a seamless flow of collaboration and could result in a performance increase. The results presented in this chapter can support this hypothesis.

In conclusion it can be said that maintaining the flow of collaboration is likely to be important in supporting group interaction and that immersive CVEs support a flow of object focussed conversation that is hard to distinguish from the real world, however object manipulation, at least without haptics and the current IPT interface, presently interferes with the flow of the task. However, users interact with objects in a spatially natural way with important aspects of their body movements represented remotely with respect to the object. Immersive CVEs appear promising for the support of distributed group collaboration and seem to be suitable for studying people's behaviour during such collaboration. Yet how much does the interface compared to collaboration contribute to the success or failure of a closely-coupled task? To answer this question some aspects have to be singled out to increase the focus and the next chapter will try to address some of these issues.

Chapter 7

7 Evaluating display properties

The previous two chapters investigated how multiple users collaborate with each other when connected from distributed sites. They demonstrated that interfaces and displays can hamper closely-coupled collaboration. This raised the question which factors in particular are responsible and this chapter aims to answer this by studying how a single user in the Virtual Gazebo benchmark is affected when using different interfaces and displays.

7.1 Hypothesis-3: display properties determine effectiveness of collaboration

The study presented in this chapter analyses a similar task (see Chapter 6) carried out by a single user, so that factors affecting collaboration can be isolated. The aim is to understand the impact of using a CAVE-like display on user-to-object interaction, so that it is possible to isolate this from previous results that showed an improvement in multi-user cooperation through shared objects. This will tell us if the advantage comes from more natural interaction with objects or more natural interaction with other participants through and around objects. Section 7.2-7.4 introduces the task and the setup for the various displays. The results are given in section 6.5 & 7.6, discussed in 7.7 in relation to previous studies and section 7.8 summarises the findings.

7.2 Task description

In order to understand how different display factors and interaction methods influence a task designed for closely-coupled collaboration the existing benchmark application was modified to allow for single user interaction. The task and task goals (Figure 6-1, reproduced from Chapter 6) remained the same (as in Chapter 6.2, Table 6-1, p. 92) with the difference that gravity was removed for selected objects to allow for manipulating those objects individually.

Objects still have to be carried to the construction site and eventually fixed with the appropriate tools and materials. For example, a beam can be inserted into a metal joiner or foot and then fixed in place by drilling a hole and fixing in a screw. The original task required teamwork, as simulated gravity required two people to lift a beam and one person to hold a joiner to a beam while it was being fixed. The need for team work was removed by disabling the simulation of gravity for beams and joints. The trial was then able to focus on single user interaction with objects. Clearly, interaction would be altered by the lack of gravity, but it was considered that the effect would be negligible.

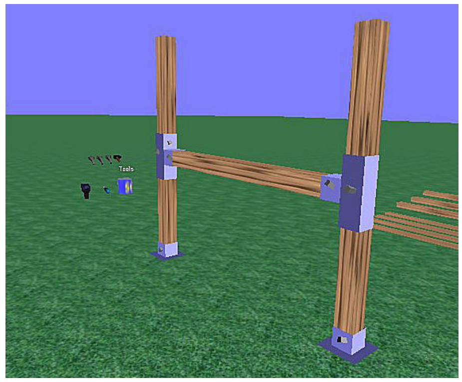


Figure 7-1: *The task target of a simple structure (vertical beams were already standing)*

7.3 Display Configuration

All participants were asked to perform this task on a variety of distinct display configurations: a non-immersive desktop system (look-into), a partial immersive workbench system (reach-into) and a fully-immersive CAVE-like (step-into) system (see Table 7-1). Each display was different in its participation frame (discussed in Chapter 3.6) that was look-into, reach-into and step-into. Each trial was first undertaken on a desktop and then repeated on the workbench and in the CAVE-like display.

Table 7-1: *Display configurations*

Display device	Input device	OS	Stereo	Field of view	Manipulation technique
Desktop (Figure 7-2)	keyboard and mouse	Linux	No	60 degree	ray-casting
Workbench (Figure 3-9)	tracked wand	Irix	Yes	110 degree	virtual hand
CAVE-like (Figure 3-8)	tracked wand	Irix	Yes	160 degree	virtual hand

Figure 7-2: *Desktop setup*

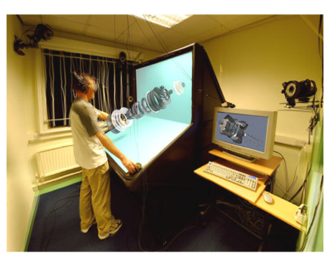


Figure 7-3: *Workbench display*

Figure 7-4: *CAVE-like display*

Tests were undertaken over a number of days. Like in previous trials (Chapter 5 and 6) the DIVE CVE was used with activated event monitoring plugin and an event filter.

An updated version of the Virtual Gazebo application was used were all object behaviours were no longer implemented as a Tcl/Tk script but as a C++ plugin. The advantage of this was a greater reliability of the application as the plugin execution was quicker and better optimised than the Tcl/Tk script. Furthermore, the plugins were modified to allow a single user to build the Virtual Gazebo structure.

Throughout the trials various measurements were taken. Tools for measurement were the above mentioned event monitor plugin (for path analysis) and time measurements of sub tasks. In addition, detailed observations were taken during the trial from an observer.

7.4 Test Conditions and Questionnaire

For this task 11 student volunteers were asked to participate, each received multiple training sessions to familiarise them with the interface and the task. Earlier trials showed that after three short training sessions the user became familiar with the interface so that their performance reached that of an expert user (Chapter 4 and 5). These trials needed no longer than 5-10 min per session and display, compared to almost 30-45 min of training and familiarisation per subject. Such a long learning process reduced earlier found problems of users not able to focus on the task, which was underlined by missing user statements regarding interface problems. In addition, to reduce variations between users, all performed the task in the same order that was first on the desktop, then workbench and last CAVE-like display.

Thirteen questions were asked, in which the user compared the different display combinations. Errors arising from a user's misinterpretation of a question were reduced by asking sets of related questions. Answers were given on a Likert-type scale [Sitzman, 2003] of 1-7, where 1 represented agreement to a very small extent and 7 to a very large extent. Those scales were later converted to percentage in order to allow for better and clearer comparison. The questionnaire included questions concerning how subjects interacted with the object in the different configurations, as well as how they perceived the interaction with the objects. The questions were similar to those asked in previous studies allowing us to compare our earlier work (Chapter 5 and 6), but were mainly related to performance, field of view and presence. The entire questionnaire is part of Appendix A.3 and summary findings are given in the next section.

7.4.1 Additional Measurements

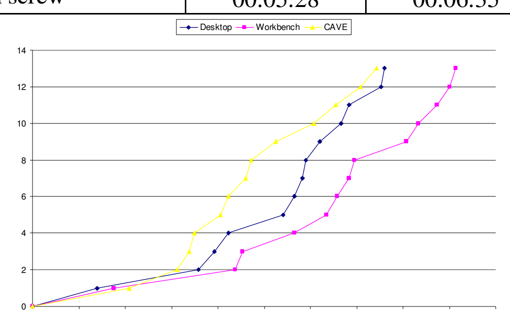
After evaluation of the results, a significant difference in measured and perceived performance was identified, this was partially related to the manipulation and navigation on the desktop. To better understand this relationship, a subsequent trial with four people was performed, repeating the desktop trial with ray-casting as well as virtual-hand manipulation (see 7.7 for discussion).

7.5 User Performance

Although the questionnaire was used to measure the user's perception of their performance, the time taken by each subject to complete a test-run was taken independently. The performance measured (Table 7-2) by time appears to contradict the subject's perception measured from the questionnaire, as discussed later in section 7.7. Average task completion times were 6.1, 7.0 and 7.3 minutes for desktop, workbench and cave respectively. An ANOVA for the measured time showed no significant difference between any of the displays.

Table 7-2: *Timing for all three display configurations*

session		Desktop	Workbench	CAVE
	example			
	Start	00:00:00	00:00:00	00:00:00
ST1	Place T joiner	00:01:00	00:01:16	00:01:30
ST2	Drill hole	00:02:35	00:03:09	00:02:15
ST3	Insert screw	00:02:50	00:03:16	00:02:26
ST4	Tighten screw	00:03:03	00:04:04	00:02:31
ST5	Place T joiner	00:03:54	00:04:34	00:02:55
ST6	Drill hole	00:04:04	00:04:44	00:03:03
ST7	Insert screw	00:04:12	00:04:55	00:03:19
ST8	Tighten screw	00:04:15	00:05:00	00:03:24
ST9	Carry beam	00:04:28	00:05:49	00:03:47
ST10	Place beam	00:04:48	00:06:00	00:04:22
ST11	Drill hole	00:04:55	00:06:17	00:04:43
ST12	Insert screw	00:05:25	00:06:29	00:05:06
ST13	Tighten screw	00:05:28	00:06:35	00:05:21



summary for all trials in min	mean	SD	mean	SD	mean	SD
ANOVA results (α=0.05) $F(2,30)=1.33$, $MS_W=4.21$, $p=0.280$	6.1	1.4	7.0	1.8	7.3	2.1

Table 7-2 also shows the timing for a typical session where slight differences between the desktop and the immersive displays are observed. At the beginning the user carried the joints to the structure, where he had to rotate and place them. This took longer on the desktop compared to the CAVE-like display as rotating an object with a 2D interface is not easy compared to a intuitive hand rotation. The user progressed on the CAVE-like display faster than on the desktop until carrying and placing the beam. This took longer on the immersive displays as the large scale of the beam (2m length) required changing perspective to properly place the beam between both t-joints. In contrast, on desktops users preferred to move the beam from a distance where positioning and keeping an overview was easier. This shows that various forms of interaction with different size objects can have an effect on the performance.

7.6 User Evaluation

This section documents the results of this study, comparing user performance, manipulation technique, FOV and presence. First the questionnaire results are presented and then the observations and measurements of two selected cases are further examined.

For the analysis of the questionnaire, similar to earlier chapters, the statistical approach of analysis of variance (ANOVA) was used to examine the extent of variations between different groups of questions. If a significant difference could be found, a posthoc Tukey test was applied to further examine the variations observed.

Note: For presentation reasons the table attachment below is not further attached to any table in this chapter yet used throughout.

Where:	α is the limit of significant deviance MS_W is the mean square within groups F(a,b) is the variance between groups / MS_W p is the actual deviance, with four decimal places M is mean & SD is standard deviation	significant difference between: [green] and [orange] as verified by the posthoc test (Tukey)

7.6.1 Overall Findings

The first question asked users *"how well they performed the task of carrying / fixing an object using the different displays"* and an ANOVA showed that there is a significant difference between the desktop and the immersive displays (performance carrying and fixing). On a desktop, performance was perceived to be less effective than it was in the CAVE or workbench. In addition, this contrast was stronger for fixing an object than for carrying it (Figure 7-5, Table 7-3).

Table 7-3: *ANOVA results for application and interface questions*

Perception of in %	Desktop		Workbench		CAVE		ANOVA results (α=0.05)
	mean	SD	mean	SD	mean	SD	
Please give your opinion of how well you **performed** the task of ...							
… carrying …	68.8	12.5	75.3	11.2	81.4	13.6	F(2,29)= 2.70, MS_W=0.756, p=0.084
… fixing …	74.3	13.1	77.9	14.8	91.4	10.0	F(2,28)= 4.96, MS_W=0.812, p=0.014
To what extend did the following **hamper** the task …							
interface	61.0	23.1	49.4	20.6	37.7	17.2	F(2,30)= 3.59, MS_W=2.05, p=0.040
delays	21.4	18.1	18.6	9.6	18.6	6.9	F(2,27)= 0.17, MS_W=0.767, p=0.841
inconsistencies	25.7	21.1	22.9	18.1	22.9	13.8	F(2,27)= 0.08, MS_W= 1.57, p= 0.919
To what extend did the following **important** …							
field of view	53.2	26.4	68.8	14.0	77.9	24.2	F(2,30)= 3.47, MS_W=2.42, p=0.044
navigation support task	64.9	21.5	70.1	19.6	74.0	19.0	F(2,30)= 0.57, MS_W= 1.98, p= 0.573
missing sense of touch	41.6	28.2	54.5	21.0	59.7	27.7	F(2,30)= 1.44, MS_W=3.27, p=0.252

The question of "*how much did the interface hamper the task*" illustrates a significant difference between desktop and immersive displays. The keyboard/mouse combination on the desktop system with its complicated combination of shortcuts (typical for CVEs) was clearly perceived to hamper the task much more than the tracking / joystick combination in the CAVE or workbench (Figure 7-5, Table 7-3). No significant difference could be found for delay or inconsistency problems. The low values are in accordance with the observations that no such problems occurred.

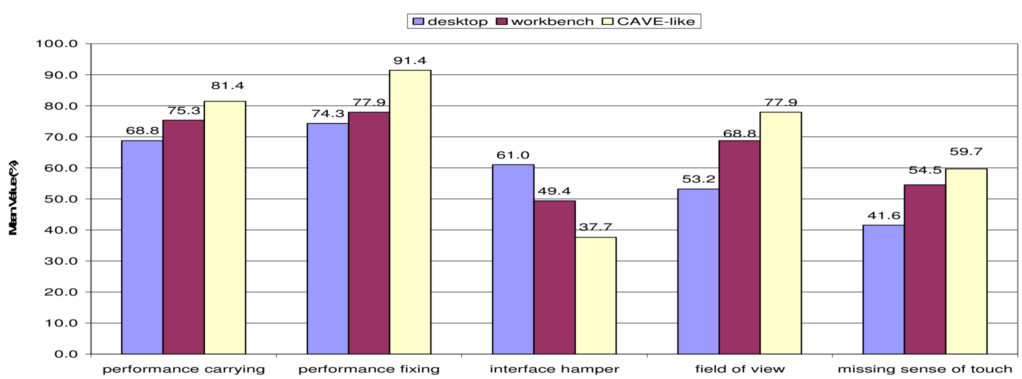


Figure 7-5: *Questionnaire overview to perception of application and interface influences*

Another question was "*how important was the field of view during the interaction*" and again a significant difference can be seen between the desktop and the immersive display (Figure 7-5, Table 7-3).

Answers to the question "*To what extent was the navigation supporting the task*" showed no significant difference, however, high values for the immersive displays indicate their usefulness (Table 7-3).

None of our displays had a haptic interface and when asking: "*how much did you miss the feel of touch*" an analysis showed that it was missed more within the immersive displays than at the desktop system. One of the users expressed it as: "The sense of touch was not expected when using the desktop, whereas it was when on the workbench and particularly in the CAVE." (Figure 7-5, Table 7-3).

Table 7-4: *ANOVA results for immersive influences*

Perception of in %	Desktop mean	Desktop SD	Workbench mean	Workbench SD	CAVE mean	CAVE SD	ANOVA results (α=0.05)
presence	28.6	18.1	66.2	9.6	85.7	14.3	F(2,30)= 44.67, MS_W= 1.02, p=0.000
physical space	32.5	18.2	71.4	18.1	89.6	12.9	F(2,30)= 34.14, MS_W= 1.35, p=0.000
interaction appear realistic	28.6	20.2	55.8	14.9	75.3	12.9	F(2,30)= 22.81, MS_W= 1.30, p=0.000
lack of social feeling	28.6	20.2	29.9	21.6	33.8	27.3	F(2,27)= 0.15, MS_W= 2.65, p= 0.862
work as a team or alone	31.4	29.2	58.6	28.1	77.1	27.9	F(2,27)= 6.54, MS_W= 3.96, p=0.005

with significant difference between: [green] and [yellow] and [orange]

The results above demonstrate that the user in the immersive display felt more natural and present in the task. This was confirmed by their answer to our questions regarding presence. The questions "*of their sense of being there", "realistic appearance of interaction*" and "*feeling of physical space*" show all a very low perception of presence on the desktop, but a high perception on the immersive displays (Figure 7-6, Table 7-4). Interestingly, a significant difference was found by ANOVA analysis between *all three* displays.

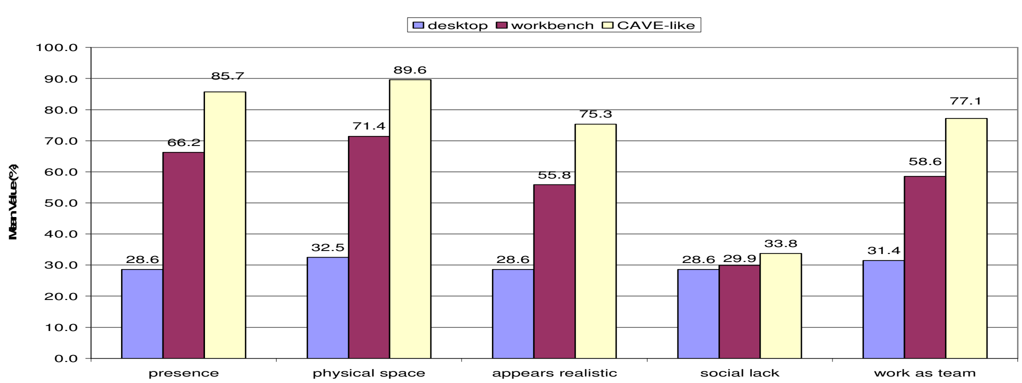


Figure 7-6: *Questionnaire overview to perception of immersive influences*

A question regarding the need for social feeling during the task did not show any significant differences and low values indicate that it was not missed (Figure 7-6, Table 7-4). In addition a hypothetical question was asked to find out "*To what extent, if at all, do you think it would have been easier to work as a team or alone*". An ANOVA analysis demonstrates a significant difference between the desktop and CAVE-like display. All other results show that the CAVE-like display was easier to use and gave a better feeling of immersing into the environment. Therefore, it could be argued that teamwork is preferred for immersive displays as this would better suit the display capabilities.

7.6.2 Comparing two extremes in Detail

The results above illustrate that the users perceived the use of immersive displays as more efficient and suitable than the use of a desktop display. However, these results contradict the task performance measurements. The average time to complete the task was similar and with no significant difference for each display. This contradiction will be discussed later in this section, but first two opposite cases will be presented (Table 7-5).

Table 7-5: *Comparison of two opposite cases, using 7-point Likert-type scale*

perception of	Desktop	Work-bench	CAVE-like
case1			
measured task time	6 min	6 min	6 min
main observations	- good use of all walls in the CAVE - "10min ago I was working on the wall, now I am in the middle and that makes a difference"		
case2			
measured task time	6 min	7 min	9 min
main observations	- a mental picture of the scene seems to be missing - stayed in one place in the immersive display, but lots of joystick movements		

In the first case (case1) the user had an equally fast performance time on all displays and in the second case (case2) the desktop time was faster than on the immersive displays. The main difference between the two has been observed in how they used the display interfaces. The former was taking advantage of the display's properties (manoeuvrability, field of view, interaction technique), whereas the latter used all displays as if he was fixed in his position (Figure 7-7, Figure 7-8, Table 7-5).

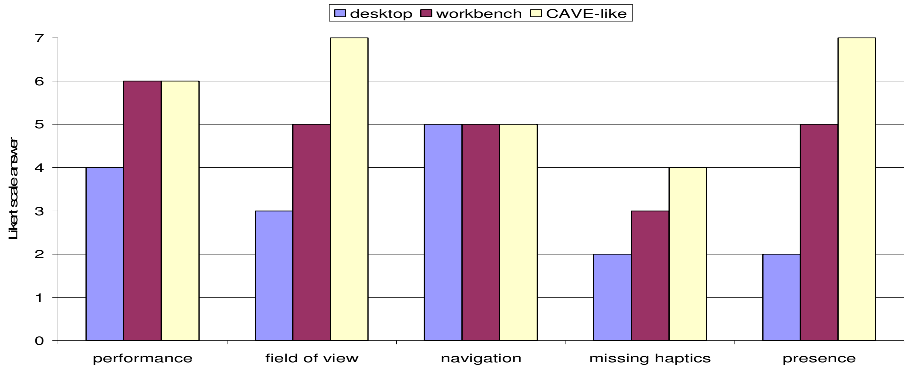


Figure 7-7: *User perception of case1 (with observed flexible interaction pattern)*

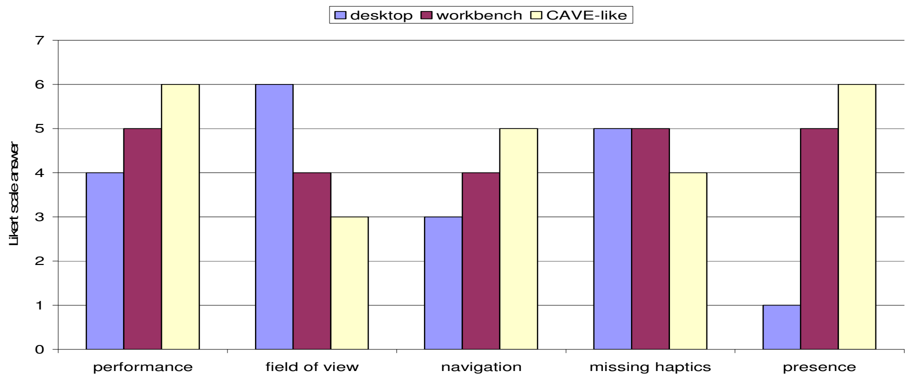


Figure 7-8: *User perception of case2 (with observed ridged interaction pattern)*

The CVE platform used in this trial allows manipulation of objects through ray-casting on the desktop display, whereas a user must physically reach for an object before it can be manipulated through the immersive displays. This has the effect that the desktop user can manipulate objects from a distance, whereas in the immersed setting they must first approach the object. The advantage on the desktop is an apparent increase in the "field of view" when the building site is viewed from a distance. However, this would only work well in an open environment, as it is the case in the experimental setting in this trial. In a normal sized room, surrounded by walls, it would be difficult to see the whole room, and subsequently this would make it necessary to turn around. The effect of a large open environment can be seen in Figure 7-9a (traces show navigation through environment) where the desktop user moved very little and performed the object manipulation from a distance.

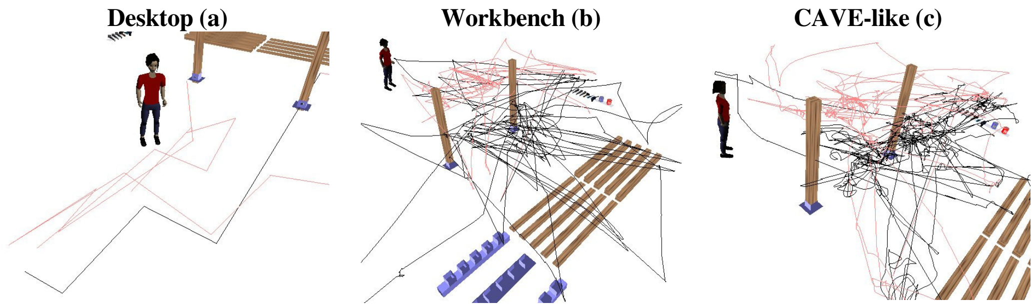


Figure 7-9: *Traces of the moving avatar during the task, case1: dark line, case2: bright line*

In contrast, in the given configuration, the immersive displays required direct manipulation, hence the large amount of movements for both users in Figure 7-9b and Figure 7-9c. In addition, a larger amount of movements in a contained space has been recorded for the CAVE-like display compared to the workbench. From observations, it could be argued that this could be due to the difference in modes of interaction across the display types. The CAVE-like display was a 3×3m room in which the user can freely walk due to the tracking of the body, allowing natural precise and fast movements around an object, if it is close enough (within the 3×3m). This includes the ability to swing the body around, using peripheral vision and eye cascades to control an effective turn to an object of interest, when displayed on another projection wall. The joystick controller is only needed for larger movements. In contrast, on the workbench the user is more restricted (space of 1×1.5m) by the physical space as well as the smaller FOV, making it necessary to use the joystick controller more often for navigation. This can be seen in comparing the fairly straight lines of Figure 7-9b (using joystick navigation) with curved lines of Figure 7-9c (user walking within the spatial display). In addition, the Figure 7-9(a-c) shows that the user of case1 is moving less and shorter than the user of case2. This is in harmony with the observations that in case2 the joystick was used more often than in case1, where the user made more use of his physical space to move. The result is an increase of measured completion time of the task for case2.

Observations have shown that taking advantage of the natural interface of the immersive display could increase the feeling of presence and performance (Table 7-5) as well as reducing the frustration factor, because one may “overshoot” the target when trying to get there with the joystick. Similar observations have been made in previous trials during closely-coupled interaction (Chapter 5 and 6), where overshooting led to some observed distress when a user needed more time to adjust their position. Thereby the other user had to wait if one’s action was needed to finish a cooperative subtask.

7.7 Discussion

This section discusses why perceived and measured performance was different, what the FOV has to do with user locomotion & navigation and why the interaction technique influences the user collaboration and performance.

One clear observation was the difference between the perceived performance and the time needed to complete the task (see Table 7-2 and Table 7-3). The contradiction of these results may be explained through the relationship of the perception of being there, immersiveness and interaction technique (Figure 7-10).

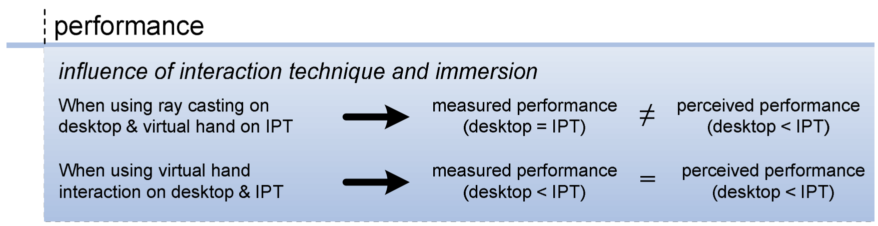


Figure 7-10: *Correlation diagram of perceived and measured performance*

The results of this study (Figure 7-6) show significant differences in perceived presence for all displays. The same tendency can be seen for performance, FOV, missing touch and interface problems. Although those tendencies are not as strong as for presence, they show that the more one is immersed and engaged the higher is the feeling of being there. Presence is not something that can be clearly measured, but is a feeling created by a number of factors (Chapter 2 and 3). Those factors, like immersiveness, naturalness of interface and ease of interaction, all appear to contribute to a feeling of being there. Small differences of perception (between displays) for all those factors have a profound influence on the perceived presence. This also explains the difference between perceived and measured performance. If one feels more engaged and present, time will seem to pass quicker and the user's own activity will enhance the feeling of performance. This can also be seen in the reaction of users, who consistently mentioned that the use of the immersive display was much more enjoyable ("*Great experience and fun!*") than the desktop.

One objective in this study was to determine how much the FOV would influence task performance. The hypothesis was that with a wider view frustum the task would become easier and increase performance as the scene is more visibly accessible and therefore objects can be spotted quicker. In contrast to the desktop, both immersive displays are similar in the way the user interacts, however the FOV is their main difference. Differences can be seen in the data and observations gathered during this trial. At the workbench, Figure 7-9b shows clearly longer paths for locomotion in comparison to Figure 7-9c. In addition, the observation during the trial was that on the workbench the joystick was used more often to attain an object as compared to the CAVE, where physical walking toward an object was easier and only longer distances needed the use of the joystick.

Orientation through rotation is a behaviour people do all the time and in virtual environments joysticks (keyboard) are used to achieve a full rotation. Exceptions are fully immersive displays like HMD or 5-sided CAVE, which have a natural rotation of 360º, independent of the FOV. This means that with an HMD the user may not need to use a joystick to rotate, but rather uses his own body [Bowman et al., 2002a]. In contrast, the desktop has the smallest FOV of all the tested displays, yet the locomotion recorded during the trial was very low. The reason for this appears to be based upon the ray-casting manipulation of objects. The user did not need to get close to the object, but could do everything from a remote place, from which the whole scene could be observed. However, in previous trials this behaviour was reason for complaint as other collaborating users could not see the correlation between a user and the object they were interacting with ([Hindmarsh et al., 2000; Roberts et al., 2004b], Chapter 5). In addition, working from a remote place is only possible if the given environment supports such behaviour, for example, a world without walls or very large rooms.

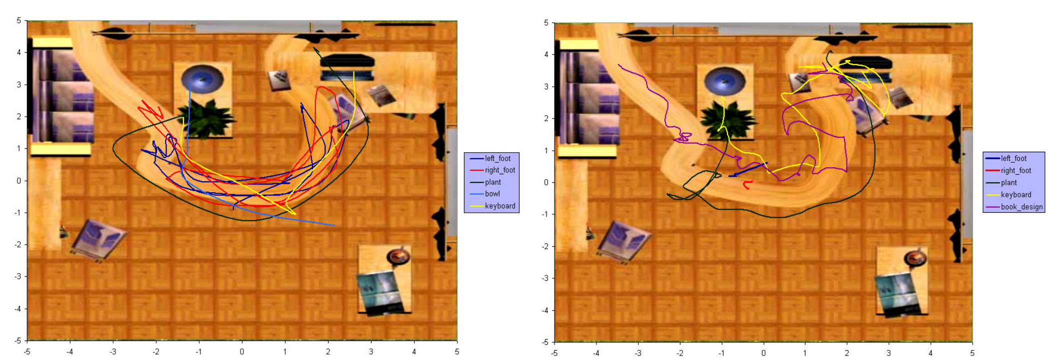
Therefore, in a subsequent trial to this study users were asked to repeat the task on the desktop, first from a several metres away in the virtual environment (using ray-casting) and then from a location adjacent to the object (virtual-hand). The result was that the time taken to perform the task doubled for the close-up trial (mean of 9.4min). Therefore it can be hypothesised that if we try to improve the collaboration between users by allowing only close-object interaction, performance time for desktop user will drop due to their limitation in FOV hence resulting in extended locomotion time to orientate (see Figure 7-10). In addition, a study from Steed et al. [2005], that compared ray casting and virtual-hand interaction on HMD and CAVE displays, found that virtual-hand is superior for selection and manipulation of objects.

This study looked into influences on a single user task. Those influences sustain in a co-presence situation and may even increase. For example, problems with interface and manipulation of objects can interrupt the workflow in a closely-coupled situation [Hindmarsh et al., 2001]. The previous studies showed that people have a higher perception of the performance of an immersed user, independent of the assessment of themselves or others (Chapter 5 and 6). They also show a significant difference between two immersed users and a desktop user, which was related to the ease of manipulation and navigation.

From the above discussion it is clear that highly immersive displays have a positive effect on users comfort und usability of the VR environment. It also indicates that immersive displays support the workflow and thereby create a performance increase. A frequently asked question is how this can be applied to other fully immersive displays like HMDs. The comparison of IPT and HMD has surprisingly not been widely studied. A lack of user studies is reported by Manek [2004] and a complete absence of direct comparisons reported by Steed et al. [2005]. The few studies available are restricted to selection, manipulation or locomotion tasks.

A related study by Sander was using a novel approach of placing a HMD in the CAVE-like display [Sander, 2005], it allowed participants perceiving the environment through either, and to be observed moving within the IPT holograph. Combined with sharing the same tracking and camera systems, this provided a direct comparison of tracking measurements, interaction behaviour, perception, video and other observational data. The experiment studied participants moving objects around a living room setting initially on a level surface and then whilst varying the height and shape of the walking surface through raised boards. Performance in the synthetic environment, using both display types, was compared to that in a physical mock up of the living room.

Using the same tracking system and space for both immersive displays has the advantage that tracking data can be matched. Figure 7-11 shows a series of such trajectory graphs where the CAVE-like user had no real problems to balance along the board, this was much harder using a HMD.



a) **b)**

Figure 7-11: *Path traces while repositioning of objects during balancing on a shaped board a) in the CAVE-like display verses b) HMD (reproduced here from [Sander, 2005])*

Observations showed that people behave quite naturally in the CAVE as opposed to HMD usage. This can also be seen by looking at the tracking data, as Figure 7-11a shows for the CAVE smother paths verses erratic paths at the HMD (Figure 7-11b). It was argued that this is due to the different properties of both displays with a lower field of view and only a partial visible virtual body in the HMD VE. Unlike the CAVE, participants using the HMD noted that they perceived a difference of the virtual and real hand position. This error in preproiception [Mine et al., 1997] was due to a limited accuracy of the used magnetic tracking system. Such problems are less notable using a CAVE as a user has his own body as a reference. Higher precision in tracking with a higher quality display may reduce such problems when using a HMD.

7.8 Summary

The measurement of performance is always difficult to achieve, as it depends on the way measures are taken and how quantifiable a task is. This applies as well for performance in a virtual environment. One may be able to measure the time it takes to finish a task, but as this study shows this is not necessarily reconciled with perceived performance.

Results from combining this chapter with the last two suggests that people always think they are performing better in a "step-into" CAVE-like display than on a "look-into" desktop or "reach-into" displays. However, objective measurements of task performance indicate such improvements in object focussed collaboration but not for single user object interaction.

The results of this chapter also clearly underline that immersive displays have a significant positive impact on the perception of presence and realism. Although chapter 2 discussed this in detail and made references to previous work with similar results, this work is contributing to the discussion from the approach of close object interaction. A frequently asked question is how this can be applied to fully immersive displays like HMDs. Placing a HMD in the CAVE-like display [Sander, 2005] allows participants to perceive the environment through either, and to be observed while moving within. It is likely that the outcome of those tests is determined by the quality of the HMD used. In this particular case a non state of the art HMD seemed to have a clear effect on people's behaviour.

The studies discussed have shown that different factors lead to an increasing perception of presence and performance. Factors such as FOV, manipulation technique and navigation, may also influence a user's interaction and its effect on other participants in a collaborative task (e.g. no fragmented workflow). In addition, different displays influence the way users interact and behave to a degree that makes working in the environment very difficult or even causes sickness. This means that a display has to be suited to a task and that the design of such a task ought to incorporate the display and interface properties. A more detailed discussion about various influences on a task concentrating on closely-coupled collaboration will be given in the next chapter.

Chapter 8

8 Factors influencing distributed collaboration in immersive CVEs

Natural object-focussed collaboration in distributed virtual environments is the main focus of this work and this chapter will discuss what is known and has been learned to improve such collaboration. So far this thesis has reasoned that immersive CVEs are suitable for distributed collaboration and various studies using the Virtual Gazebo benchmark application have proven that such collaboration is possible and can be effective. Chapter 1 introduced the notion that the research problem for natural object focussed collaboration is more than just connecting immersive displays using CVEs (see Figure 1-1, pp. 6) as it depends on the right settings.

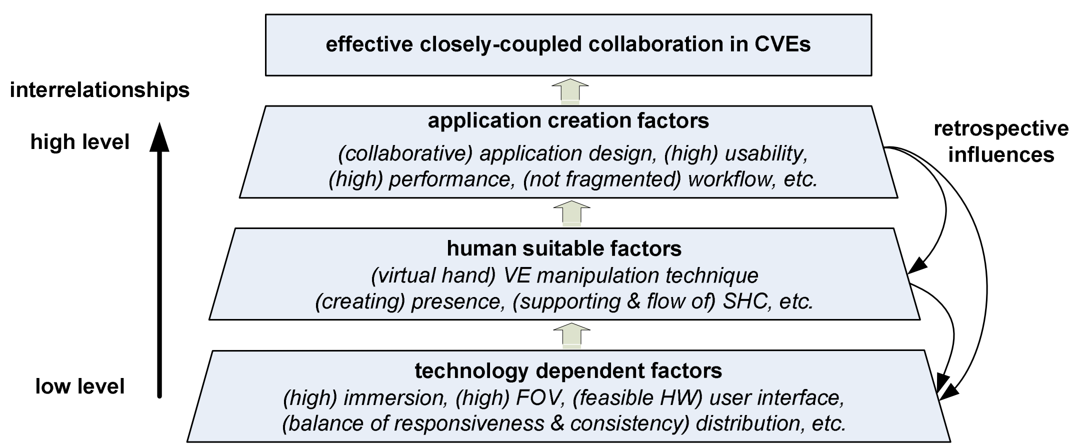


Figure 8-1: *Different factors that can influence collaboration*

The choice of display and technology affects immersion, field of view (FOV), interface and these subsequently workflow, social human communication (SHC) and performance. In addition, the choice or respectively design of application needs to reflect display properties, and influences through distribution and task design. For example, the user interface does not necessarily depend on workflow, but the workflow depends on a usable user interface, application design, support for SHC, etc. Therefore we can categorise those factors into different relationship levels supporting closely-coupled collaboration (Figure 8-1). At the lowest level we find technological factors such as immersion, FOV, user interface, distribution, etc, which influence human factors (e.g presence, SHC, manipulation technique, etc). Both levels have further impact on application factors such as task design, usability, performance, workflow, etc. However, the higher levels have impact onto the lower levels as well.

This research has looked closely into some of these issues whereas others are well known and discussed in the early part of this thesis. Like in a big puzzle these factors have to be combined to understand the bigger picture of its multi-dimensional relationships. Therefore, the following sections will discuss the above interrelationships in more detail and will combine various parts of the thesis to build on overall picture (Figure 8-1). The discussion is starting at the higher levels and shows how they relate to closely-coupled collaboration using immersive CVEs combining various parts of the thesis to build an overall picture.

8.1 Application creation factors

This thesis postulates that straightforwardness of interaction with others through and around objects will be of value in many collaborative scenarios. This section discusses major application factors (Figure 8-1) that impact on the straightforwardness.

8.1.1 Workflow

Disruption in the work process from re-orientation within the virtual scene can have implications when working together, because teamwork relies on a fluent and coherent multimodal communication between the partners that can be disrupted by the technology. For example, a person may reference an object by speaking its name and pointing to it, but technology may fragment the workspace such that the meaning of the gesture is lost or it may fragment the nuance such that the gesture occurs at a different time to the spoken description. It usually means that the partner needs to re-orientate and interrupt his work in order to see each other and the object they are working with [Hindmarsh et al., 2000]. This fragmentation of the workflow can be time-consuming and therefore could have a negative impact on the performance. Figure 8-2 summarises some of the complex relationships and dependencies to support a seamless workflow.

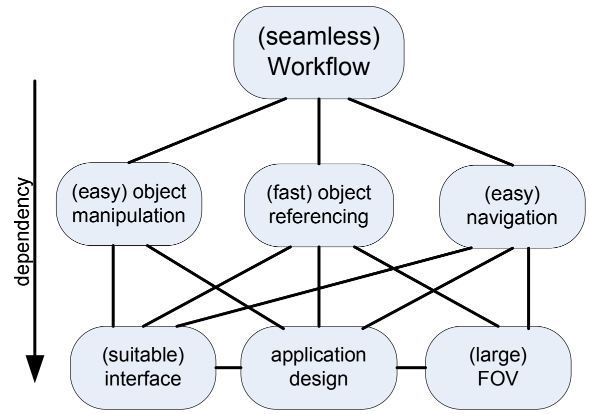


Figure 8-2: *Dependencies for a seamless workflow*

In the study by Hindmarsh et al. the authors found that the limited FOV on desktop systems was of great hindrance due to problems with fragmentation of the workspace and they concluded that this was caused from a lack of information about other's actions due to their limited "window into the world". Similar issues could be found in the CAVE-desktop trial of Chapter 5 where participants had difficulties to see which desktop user is manipulating a certain object. A subsequent study by Fraser et al. tried to resolve some of the issues with peripheral lenses, which resulted in an enhanced FOV. Although this solution enhanced the awareness it showed that peripheral lens distortion can disrupt both a user's own sense, and their notion of the other's sense, of orientation to actions and features within the environment [Fraser et al., 1999]. In contrast, using an immersive setup to "step-into" the environment (Chapter 6) does not seem to hamper the workflow by such an extent as users directly interact with objects (Table 8-1).

Table 8-1: *Influences of manipulation technique during collaboration on workflow*

remote manipulation using ray casting	→	fragmentation of workflow
close-by (virtual hand) manipulation	→	visible connection between object and user, no fragmentation

The natural use of the body in IPTs to reference and interact with objects can increase both task performance and subjective impression of closely-coupled collaboration (Chapter 5 and 6) and an initial trial indicates that the scale of this improvement is relative to the spatial extent of the task [Roberts et al., 2005a]. Results indicate that communicative gaze plays a strong role in the performance by demonstrating focus of attention. As the physical extent of the display is considerably less than that of the shared environment, this advantage is unlikely to be connected to the mode of navigation. A more likely contributing factor is the reduction in fragmentation of the workflow, brought about by bringing two people within the same shared space and allowing each to see where the other is looking and pointing from a natural perspective. Table 3-1 in chapter 3 summarised and compared, among other things, various technologies and their support for distributed collaboration without interrupting the workflow.

8.1.2 Task Performance

The goal of developing new technologies is usually to improve efficiency and performance of an existing task. Task performance can be measured in many ways. Common are the quality of output and the time it took to achieve. Chapter 5 measured an increase of performance in a collaborative task in CAVE-like displays compared to desktop displays, yet no such difference could be measured on a single user task in Chapter 7. At the same time both showed an increase of perceived performance. These studies have shown that measured performance does not necessarily reconcile with perceived performance (Table 8-2). However, objective measurements of task performance indicate such improvements in object-focussed collaboration but not for single user object interaction.

Table 8-2: *Influences of manipulation technique and immersion on performance*

When using ray casting on desktop & virtual hand on IPT	→	measured performance (desktop = IPT)	≠	perceived performance (desktop < IPT)
When using virtual hand interaction on desktop & IPT	→	measured performance (desktop < IPT)	=	perceived performance (desktop < IPT)

It has to be noted that for those trials each display used their own standard manipulation technique. However, using a manipulation technique that supports a continuous workflow and better SHC during closely-coupled collaboration has a negative performance effect for desktop displays often doubling their time to finish the task (Chapter 7). If we look at the display alone, it appears that IPTs are better at representing activity and contribution of others as they naturally reproduce the way in which people look at, move into and reach for objects. At the same time IPTs can trick a single user into thinking they are achieving more than they truly are. However, effective closely-coupled collaboration is not depending on only one factor but many and each contributes to increase performance (Figure 8-3).

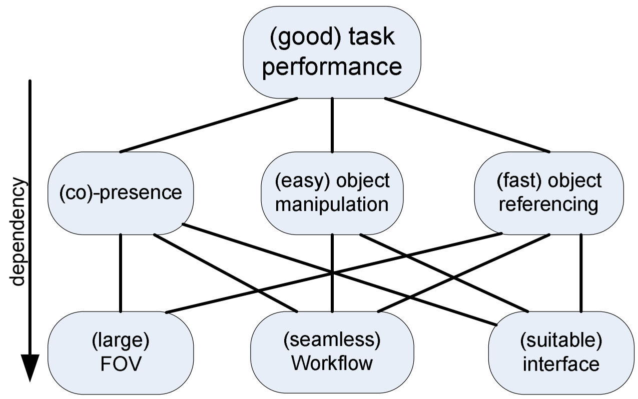


Figure 8-3: *Dependencies for a good task performance*

Influences through immersion, naturalness of interface and ease of interaction, available FOV, as well as manipulation technique all appear to enhance performance in collaborative scenarios. Some of these factors are discussed further below.

8.1.3 Application Design

At present there are noticeable differences between reality and virtuality. In some applications this is not a problem, in others it is. In the latter case it is necessary to compensate for the difference, either to make the difference unperceivable or unimportant. Concurrent object manipulation by multiple users requires a design which is task supportive and provides equal feedback of actions. Sometimes a compromise has to be taken between natural interaction and restrictions imposed by environment and technology (Figure 8-4). Let us consider a case where technology characteristics change the way things work:

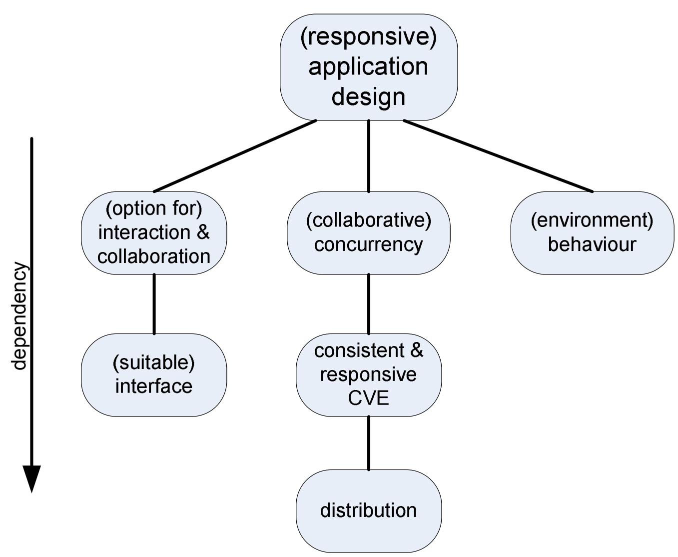


Figure 8-4: *Dependencies for on application design to support collaboration and interaction*

An effective collaborative application has additional affordances to single user applications, as actions need to be distributed and user's VR interface has to allow collaborative remote interaction. In order to work efficiently, the collaborative task needs to support human communication and appropriate ways to interact with the environment. Both are only achievable with a user interface suitable to the interaction task. For example, objects in a virtual environment could signal if and how a user is interacting with it. Usually users of a CVE can see that an object is moving, but seeing that it is selected is not always obvious. This is quite important when two or more people try to manipulate an object concurrently. Depending on the available user interface the object could show visual or even haptic cues that a certain part is selected by a user. The Virtual Gazebo benchmark used visual identification to show which part of a beam was selected by a user to allow a second user to select the other end of the beam allowing both users to carry it concurrently (Chapter 4). This kind of feedback is something that needs to be considered while designing the application as it could have profound impact on the application structure. Late stage implementation may need a complete redesign of the application.

8.2 Human suitable factors

Seamless distributed collaboration with and through objects can only be achieved if people using the application and performing the work, don't have to think about it but behave as they are in the real world. This section looks closer into human factors (Figure 8-1) that impact on the straightforwardness of closely-coupled collaboration.

8.2.1 (Co)-Presence

The results of two studies showed significant differences in perceived presence for all displays (Chapter 5 and 7) and high perception for immersive CAVE-like displays (Chapter 6). Presence is a feeling created by a number of factors, and thus difficult to measure (Chapter 2, [Slater, 1999]). These factors (Figure 8-5), such as immersion, naturalness of interface and ease of interaction, all appear to contribute to a feeling of "being there". Small differences of perception (between displays) for all those factors have a profound influence on the perceived presence [Snow, 1996]. This also explains the difference between perceived and measured performance. If one feels more engaged and present, time will seem to pass quicker and the user's own activity will enhance the feeling of performance. For example, users of the CAVE-desktop trial consistently mentioned (Chapter 5) that the use of the immersive display was much more enjoyable than the desktop and that time seemed to pass quicker. Furthermore, a survey by Youngblut reflect on a relationship between task performance and presence [Youngblut, 2003] and found that there were a total of 50 findings to review and half of these showed significant correlations between task performance measure and presence. The technology overview in chapter 3 (Table 3-1, pp. 55) is also highlighting its support for creating a feeling of presence. A break in presence can result in an interruption of the workflow with negative consequences on performance [Brogni et al., 2003a]. In addition, some consider the experience of presence as a precondition for co-presence [Slater et al., 2000a] which in turn is a precondition for collaboration [Tromp et al., 1998]. Improving presence therefore improves collaboration. Chapter 6 demonstrated that co-presence was perceived very high when two immersive users interact and interestingly this was independent of difficulties encountered with the provided interface.

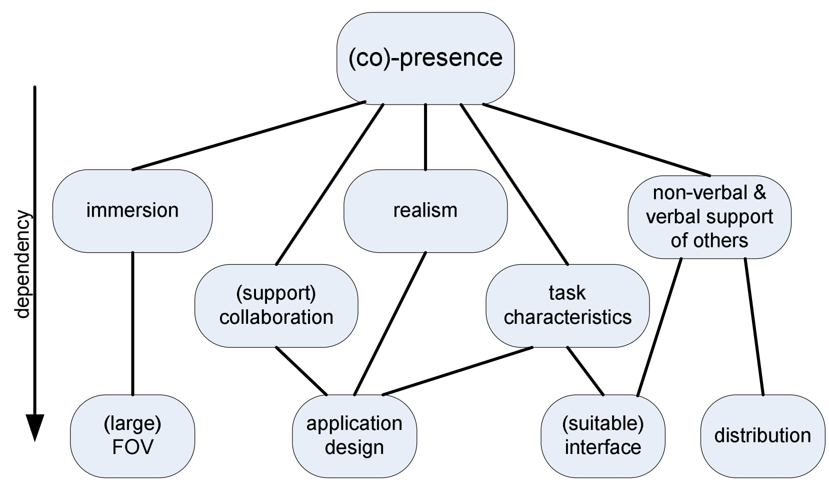


Figure 8-5: *Dependencies for supporting and creating (co)-presence*

8.2.2 Social Human Communication

Another reason for interruptions in the team workflow is insufficient support of social human communication (SHC). Communication is essential for teamwork as it makes us aware of what’s happening around us and expresses our action (Chapter 2), especially if this involves close coupling of object interaction between team members where everyone needs to understand the current work process and task goals. Supporting this accordingly is important as the flow of conversation differs between a co-located and computer supported distributed group [Riva, 1999].

IPTs physically place the user in an intuitively interactive information context and linking such displays with a CVE additionally situates people in an intuitively social context [Roberts, 2003]. This technology lends itself well to SHC by supporting its four primary elements: verbal and non-verbal communication, references to objects and references to the environment [Knapp, 1978; Burgoon et al., 1994], which were discussed in detail in Chapter 2. An overview of various communication mediums and their support for the four components of SHC as well as their support of naturalness, presence and workflow can be found in Chapter 3 (Table 3-1, pp. 55). In the real world, people combine all these elements, often without realising it. Verbal communication includes mainly the human linguistic system and its derivatives (speech, writing, sign language, etc.), whereas body language, including posture and gesture, belong to non-verbal communication. Verbal and non-verbal communication are often inextricably linked through nuances such as lip-synch, clapping and unintentional gesturing and posture changes while speaking. All this is natural to people and collaboration usually incorporate these to achieve a continuing workflow. Figure 8-6 shows the main correlation between factors to support SHC for distributed groups. Supporting its primary elements as well as creating a feeling of trust, (co)-presence is essential for effective communication. In turn this has to be provided by technology and application design.

The subject of communication is not always abstract and often relates to our surroundings and artefacts within it, both providing a context for understanding. People may discuss their surroundings or an object through both verbal and non-verbal communication but in addition they can move around the environment and manipulate objects within it. A nuance might arise from the synchronisation of concurrent elements of SHC. For example, a user might point to an object saying “lets pick that up” and then turn and point to a place in the environments saying “and take it over there”, thus relating verbal and non-verbal communication in relation to an object and the environment. Such behaviour could often be observed when users connect through immersive displays (Chapter 6) even though SHC support was limited.

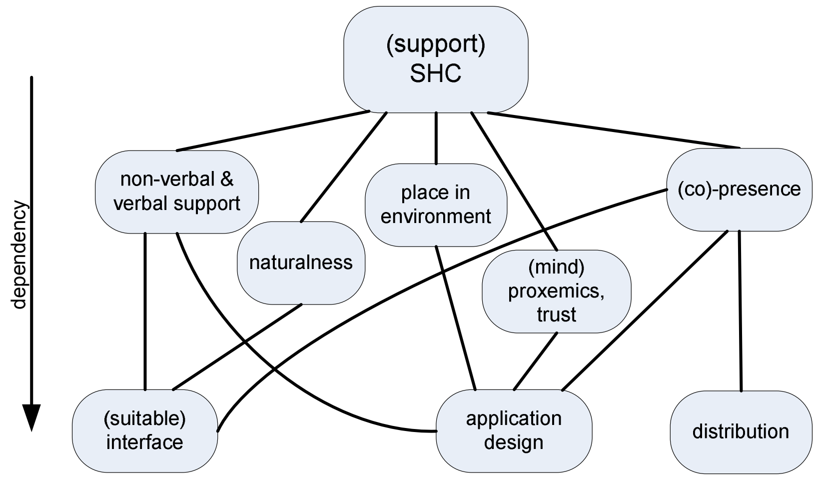


Figure 8-6: *Dependencies for supporting SHC*

Collaborative environments generally allow all participants to directly modify objects and to observe the effects of changes made by others. Chapter 2 discussed among other things that early studies of collaborative environments showed that adding this type of (non-verbal) visual information improves the efficiency of speech communication [Bly, 1988; Whittaker et al., 1993]. For example, Whittaker et al. compared speech only communication with speech complementing a collaborative environment for three different tasks: brainstorming, spatial design and collaborative editing. They showed that the environment improved communication for the latter two tasks, but not for brainstorming. Analyses of linguistic behaviour showed the reasons why: when the task requires reference to a complex layout (design and editing) or complex visual objects (spatial design), people were able to express complex spatial relations ("put that over here") and to use directed gesture. Participants were also more implicit in their communications when using the environment, because the environment supported situational awareness [Endsley, 1995]. Therefore, participants did not explicitly need to communicate changes about the current task if the collaborators could see this information directly. These effects were not found in the brainstorming task, which did not demand reference to complex objects, spatial relations or object transformations, something very common during closely-coupled collaboration and possible cause of workflow fragmentation (see Chapter 5, [Hindmarsh et al., 2000]). Initial research presented in chapter 6 suggests that IPTs appear to overcome these problems and that SHC of two immersed users contributes to a seamless workflow.

8.2.3 Manipulation Technique

Manipulating objects in desktop CVEs is difficult when using only mouse & keyboard while in addition a team needs to see what each user is doing. Therefore, a technique is required that allows straightforward manipulation of objects while supporting group interaction.

Manipulation technique is the method by which the user performs an operation on a virtual object via the interface and may be as simple as clicking the mouse button, or as complex as a series of gestures and speech. There may be many possible manipulation techniques for any given interaction task and each is effecting usability as well as collaborative interaction. The manipulation technique can be influenced by the input device used, but is not completely constrained by it. The same input device may be used for many manipulation techniques for the same task; conversely, it may be possible to implement a given technique using several different input devices.

Desktop systems use various methods to interact with objects in a virtual environment, such as go-go, ray casting or occlusion techniques [Poupyrev et al., 1998; Bowman et al., 2001]. These can be used in IPTs, but have been primarily developed using head-mounted displays (HMD). Desktop systems use 2D interface controls or mouse picking, whereas immersive displays normally use one or two handed direct manipulation (virtual-hand) using a tracking system. Evaluations of interaction techniques for immersive displays found that the virtual-hand is superior to ray casting for the selection and manipulation of objects [Poupyrev et al., 1998; Steed et al., 2005]. Experiences with the Virtual Gazebo application also showed that virtual hand interaction greatly enhances interaction and collaboration (Chapter 6). In a different study by Kjeldskov et al. it was found that partial and fully immersive displays have different support for close-by interaction (virtual-hand) and different affordances for pointing (virtual beam) [Kjeldskov, 2001]. Furthermore, in a subsequent trial to the study in chapter 7.7 the authors asked users to repeat the task on the desktop, first from a far distant location (using ray-casting) and second from a location close to the object (virtual-hand). The result was that the time taken to perform the task doubled for the close-up trial. Therefore it can be hypothesised that if we try to improve the collaboration between users by allowing only close-object interaction, time-performance for desktop user will drop due to their limitation in FOV hence resulting in extended locomotion time to orientate (Figure 8-7).

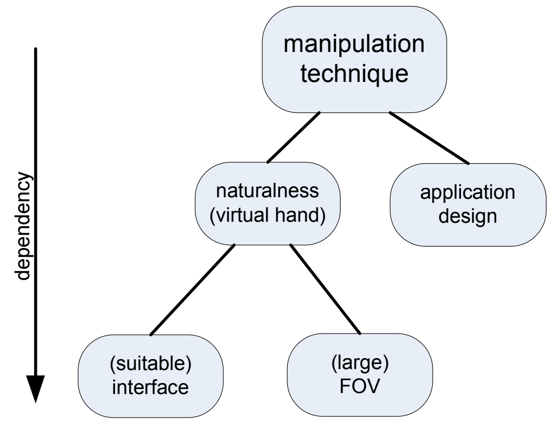


Figure 8-7: *Dependencies for choosing the right manipulation technique*

Although virtual-hand interaction on desktops requires more time for object interaction of the user, it also means that other participants of the shared environment can directly see the interaction. On immersive displays this is enhanced by body posture and even facial expressions. During closely-coupled collaboration this is very important as it allows collaborators to interact closely while maintaining a natural workflow.

8.3 Technology dependent factors

For distributed collaboration it is necessary to use technology while its properties can determine how effective and useful this collaboration can be. This section looks closer into such technology factors (Figure 8-1), their interrelationships to application and human factors.

8.3.1 User Interface

An hardware interface supporting closely-coupled collaboration has varying affordances such as support for virtual-hand interaction, appropriate navigation, transferring SHC onto a user’s avatar and giving user feedback from interaction with the VE. While in IPTs data gloves are good for object interaction it is difficult to navigate with them and joysticks are usually used instead. Our years of experience show that none of the many input devices used for the Virtual Gazebo application is fully satisfactory for the various interaction tasks during closely-coupled collaboration. Cables and complicated button allocation is one reason as well as a missing “one fits all” devices which supports natural hand manipulation and navigation.

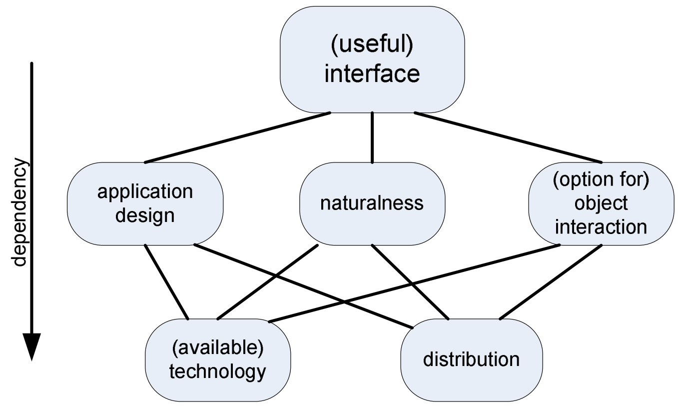


Figure 8-8: *Dependencies for choosing a useful interface*

A user interface is the hardware and software that mediate the interaction between humans and computers. It technologically includes input and output devices, such as mice, keyboards, tracking systems, data gloves, monitors, and speakers, as well as software entities such as menus, windows, toolbars, etc [Hix & Hartson, 1993]. Kjeldskov et al. found that non-tracked 3D interaction devices (Spacemouse, 3D joystick, etc.) work well for orientating and moving when using semi-immersive displays, but are problematic when using fully immersive displays [Kjeldskov, 2001]. This dependency on other factors makes it so important when choosing an interface (Figure 8-8). For example, the collaborative task of the Virtual Gazebo requires easy manipulation of small objects (tools) and efficient locomotion within the environment. The available technology during the user trials of chapters 4-7 required the user to use a tracked joypad to manipulate objects and move within the VE. This has implications on the naturalness of interaction and collaboration.

Current IPTs usually create reasonable audio-visual feedback, yet during object interaction with VE's, an important human sense is missing. The reproduction of the feeling of touch is attempted via haptic interfaces and results have been promising [Massie, 1993; Van der Linde et al., 2002]. Various interfaces have been used in applications and experiments ranging from rehabilitation [Amirabdollahian, 2004; Loureiro et al., 2005] to transatlantic interaction [Kim et al., 2004]. These interfaces, however, have been designed for desktops or other single display systems. Haptic devices often have a big footprint requiring a large workspace. Many of these mechanisms are grounded (immobile) creating force rather than just tactile feedback [Burdea, 2000]. This makes it difficult to access them within the space of the immersive display without hindering/covering the view or to reduce free movements within the display. An attempt to add haptic to the Virtual Gazebo application in the CAVE-like display demonstrated that objects could be touched but the very limited range of the haptics device made it impracticable to use [Seelig et al., 2004]. Some approaches try to avoid such problems, for example, by using exo-skeletons which are not stationary but still give us the feel of touch [Stone, 2000] or by special designed Step-in-place platforms providing free movements [Bouguila et al., 2002; *The VirtuSphere*, 2004]. However, there is still a long way to go before such devices allow natural interaction with objects and environment.

Results of chapters 4-7, gathered during trials for this thesis, suggest that people have to adapt to a new interface as well as application and that for some applications, after only a few repetitions they can achieve a performance which is comparable to expert users. If interaction in these trials was hampered, it came mainly from difficulties while grasping objects, which in turn had different causes such as loosing 3D vision when shutter-glasses lost the infrared signal, using the wrong button on the joypad or misjudging the distance to an object.

8.3.2 Immersion, Field of View and Navigation

Various display types can be used for VR, many are described in Chapter 3, and each has different properties with effects on immersion, navigation, communication and other factors relevant for supporting and improving teamwork. The term "immersion" describes the extent to which a given technology replaces real world stimuli with synthetic stimuli within the virtuality continuum [Milgram et al., 1994]. A necessary condition is Ellis' notion of a Virtual Environment (VE) as a communication media [Ellis, 1996], maintained in at least one sensory modality (typically the visual). The degree of immersion is increased by increasing the field of view (FOV) [Arthur, 1996; Lapointe & Vinson, 2002; Tan et al., 2003], greater degree of body tracking, decreased lag between body movements and resulting changes in sensory data amongst others [Pausch et al., 1997; Sheridan, 2000; Baños et al., 2004]. Immersion may lead to a sense of presence and some consider it a precondition. As discussed earlier, some consider the experience of presence as a precondition for co-presence [Slater et al., 2000a] which in turn is a precondition for collaboration [Tromp et al., 1998].

A tool often used to immerse into a VE and collaborate via CVEs is the head mounted display (HMD), they are widely used as they are more affordable and have less space constrains compared to IPTs. They fully immerse a user and come in a wide range of resolutions and different FOV. However, cost limitation and current technical boundaries allow only a limited FOV. A lower FOV results in "tunnel vision" and might decrease immersion. This can affect the flow of collaboration as such a user is less aware of its surroundings. In addition, there are ergonomic issues related to HMDs such as display's size, weight and the ability to adjust various visual parameters [Bowman et al., 2002a], slow update rates in response to head movements which can result in motion sickness, problems in proprioception (as one cannot see his own body) and HMDs don't allow to share the same display with co-located participants unless they also use HMDs. A recent study here in Salford compared an IPT with an HMD and found that with HMDs, participants "use unnatural" body posture while interacting closely with objects (Chapter 7.7, [Sander, 2005]). All these issues can affect the performance during a collaborative task as a user may struggle more with the technology rather than concentrating on the collaboration. In contrast, CAVE-like displays avoid these issues which makes it easier to interact with and through objects as well as with others.

Table 8-3: *Correlations and influences between Field of View (FOV), immersion and navigation*

low FOV	- low immersion - higher degree of locomotion for orientation
high FOV	- perceived increase of immersion - more natural / intuitive / directed locomotion

Pausch at al. found in a search pattern task that desktop users compared to an immersed HMD user took 41% more time, re-examining areas they had already searched [Pausch et al., 1997]. An objective of the single user study in Chapter 7 was to determine how much the FOV would influence task performance, comparing a closely-coupled task on three different displays (desktop, workbench, CAVE-like). A hypothesis was that with a larger FOV undertaking the task would become easier and increase performance, because the scene is more visibly accessible and therefore objects can be spotted faster. In contrast to the desktop, both immersive displays are similar in the way the user interacts, but the CAVE-like display has a larger FOV (170°) than the workbench (120°). Observations and measurements showed longer ways and joystick usage for locomotion at the workbench in comparison to the CAVE-like display (see Chapter 7.6.2).

The smaller the FOV the more locomotion can be observed, while working in a cluttered environment or closely with objects as well as people. An exception is the HMD and surrounding display's (e.g. 5-sided CAVE), which allows for natural 360º rotation, independent of the FOV. This means that with an HMD the user may not need to use a joystick to rotate, but rather uses his own body. An experiment by Bowman et al. [2002a] showed that HMD users are significantly more likely to use natural rotation in a VE than CAVE-like display users, when the CAVE-like display has a missing wall. This produces higher levels of spatial orientation, which can make navigation more efficient and improves the seamless flow of collaboration. Immersive users from this thesis trials were often observed to naturally rotate and sometimes difficulties occurred with the open wall as users had to use the joypad for reorientation, but with experience such problems can be avoided. In contrast, during the single user trial of chapter 7, the desktop had the smallest FOV of all tested displays, yet the locomotion recorded during the trial was very low. The reason for this appears to be based upon the ray-casting manipulation of objects. The user did not need to get close to the object, but could do everything from a distance, from which the whole scene could be observed. However, in the CAVE-desktop study of chapter 5 this behaviour was reason for complaint as other collaborating users could not see the correlation between a user and the object they were interacting with (Chapter 5, [Hindmarsh et al., 2000]). In addition, working from a distance is only possible if the given environment supports such behaviour, for example - a world without walls or very large rooms.

This thesis defined closely-coupled collaboration as a close coupling between object manipulation and human interaction, whereas the action of collaborating people is directly depending on each other. This includes that people can see each other as well as their motion though the environment. An immersive display with a large FOV and easy navigation would help such close interaction, as no time is spend on repositioning and orientation, something discussed earlier as reduction in fragmentation of the workflow.

8.3.3 Distribution

For enabling collaboration between remote people, all the interactions captured through an interface needs to be distributed to each participant. Within closely coupled collaboration, consistency of the shared environment, as well as responsiveness to interactions, both locally and remotely, become particularly important. Traditional CVEs were designed for desktops with only keyboard and mouse interaction. This usually is generating a low bandwidth data stream and thereby very scalable for thousands of users. IPTs on the other hand generate a high amount of tracking data (see Chapter 4 Virtual Gazebo development) requiring a high update rate for a smooth representation of the interactions by remote participants through their synthetic embodiments (avatar). A study of the Virtual Gazebo characteristics of the data stream communicated between the remote sites has shown that motion tracking causes a continuous stream of motion update events, which has significant impact on the performance of the consistency control mechanisms of the CVE software system [Wolff et al., 2004]. The Virtual Gazebo also has demonstrated that supporting closely-coupled collaboration between three immersed participants can induce delays in representing interactions of up to 3 seconds. Such delays can interrupt the workflow and concurrent manipulation of objects. Hence, the ultimate goal in the design of an immersive CVE software system is to find a balance between responsiveness and consistency [Roberts, 2003].

Two distribution models are common: server-client and peer-to-peer. The former is more prevalent in the commercial world whereas the latter is preferred in research where performance is of greater significance than security. While server based systems provide high consistency, they lack scalability and high responsiveness. This is because every client has to communicate with the server, which is in turn streaming the updates to all clients. Hence, most CVEs make use of a peer-to-peer based distribution model. Responsiveness is increased through localisation of single user interactions with a replicated database but additional consistency control is required for shared object manipulation [Roberts & Wolff, 2004a]. In addition, peer-to-peer networks are more capable of handling large amount of tracking data, thus making them more scalable compared to server-client networking, where the server would be a bottleneck.

SHC is well supported by avatars driven from live tracking data, but is adversely affected by delays, which can be induced by the network and can be proportional to the amount of tracking data sent. This is particularly the case for closely-coupled collaboration, especially concurrent manipulation of objects that require a careful balance of responsiveness and consistency in the face of the need to communicate a range of causally related SHC cues. Possible solutions include reduction of tracking data to a level just out of human perception as well as reducing redundant updates from movements of people and objects via filter or prediction ([Jung et al., 2000; Purbrick & Greenhalgh, 2003], Chapter 4). Using those solutions would enhance to transmit an increasing number of communication cues and nuances while maintaining a responsive and consistent application. Other solutions include various approaches to virtual world subdivisions [Roberts, 2003]. In addition, advanced event handling may be used to find an optimal balance between naturalness and consistency based on the actual requirements of the current collaborative scenario. For example, using multiple event-handling pipes for vital (e.g. an picking-up event) and non-vital (e.g. continues moving) allows the processing of particular event types with specific levels of consistency [Roberts et al., 2004a].

8.4 Summary

The choice of display and technology affects already immersion, FOV, interface and subsequently workflow, SHC and performance. In addition, the choice or respectively design of application needs to reflect display properties, and influences through distribution and task design. For example, the user interface does not depend on workflow, but the workflow depends on user interface, task design, support for SHC, etc. Therefore we can categorise those factors into different relationship levels supporting closely-coupled collaboration (Figure 8-1):

- low level: technology factors (immersion, FOV, user interface, distribution, etc);
- mid level: human factors (presence, SHC, manipulation technique, etc);
- high level: application factors (task design, usability, performance, workflow, etc)

It is not enough to design a system which focuses only on few of these factors, as they have influence on others with further impact on the application's usability. A system incorporating those parameters will not only have a good support for closely-coupled collaboration but also any other ("loose") forms of collaboration. Immersive CVEs appear promising for the support of distributed group collaboration and seem to be good for studying people's behaviour during such collaboration.

Chapter 9

9 Conclusion and Future Research

Demonstrating closely-coupled collaboration, as done in this thesis, is an important step towards an application allowing for many human interactions. Chapter 1 defined closely-coupled collaboration as a close coupling between object manipulation and human interaction, whereas the action of collaborating people is directly depending on each other. Margery [1999] categorised it as level 1 - co-existence and shared-perception; level 2 - individual modification of the scene; and level 3 - simultaneous interactions with an object. This work extended and clarified level 3 by highlighting the distinction between sequential and concurrent sharing of the same and different object attributes. Before natural and collaborative object manipulation can be applied in practical applications, developers and system designers need to understand how to support effective closely-coupled collaboration. The answer is not trivial and depends on a multitude of factors, a number of which were examined in this thesis.

While people cooperate with other people through an object, they use a variety of communicational resources to demonstrate their opinions, intention and needs to others. Verbal and non-verbal communication is often mutually supportive and the meaning of one can be changed or even lost without the other. This may be manifested, for example, as gesturing and posture to reinforce the emotion of the spoken word, or by talk and gesture guiding collaboration during shared object manipulation. When interacting remotely, supporting these forms of social human communication (SHC) can improve the effectiveness of collaboration. Chapter 2 introduced and discussed primary elements of SHC as well as measurement and need for a feeling of presence and co-presence to support distributed collaboration. It presented that the naturalness of collaboration depends on how well the object manipulation and forms of SHC are supported and mediated through tele-collaboration technology. However, as a survey in chapter 3 discussed, most technologies have difficulties in supporting natural object interaction with distributed groups while immersive systems seam currently to be best suited for such tasks. The survey suggests that, as of today, a system that allows users to share a common virtual space and to "step-into each others world" (Chapter 3, Figure 3-12c), such as an immersive CVE, provides the closest resemblance of co-location from distributed sites. In a CVE, remote people and shared objects can be situated in a shared synthetic environment, in which one can navigate around and interact with a computer-generated representation of objects and other participants. Thus, whereas video-conferencing systems allow people to look into each other's space, CVEs allow people and data to be situated in a shared spatial and social context.

To demonstrate and evaluate distributed collaboration the Virtual Gazebo benchmark application was created. This contains a structured task that requires at least two users to collaborate closely with each other. Communication in this virtual environment is just as important for success as technological support for human interaction and collaboration. The benchmark's successes and failures during development were documented in chapter 4 and the following chapters 5 to 7 examined the benchmark application in various user trials. This work demonstrated, within the confines of the application, that:

- distributed closely-coupled collaboration (including object focussed non-verbal communication and concurrent shared manipulation of objects) is possible with today's CVE technology (Chapter 5&6)
- a difference in level of immersion between users of linked displays leads to the more immersed user taking a dominant role, as found by Steed et al. [1999], held true for closely coupled collaboration (Chapter 5)
- however, the scale of impact depended on the closeness of collaboration, for example the method of shared object manipulation (Chapter 5)
- CVEs connected with immersive CAVE-like displays are well suited to support closely-coupled tasks (Chapter 6)
- using immersive displays supports a greater level of fluency in workflow (Chapter 6)
- People believe they are performing better in a "step-into" CAVE-like display than on a "look-into" desktop or "reach-into" displays. However, objective measurements of task performance indicate such improvements in object focussed collaboration but not for single user object interaction (Chapter 7)
- Although immersion seems to improve the perception and performance of collaboration, there are a number of factors that distract from the experience and still create a significant gap between the co-located and remote meeting. (Chapter 7&8), these and other supportive factors were summarised in a framework in Chapter 8 (Figure 8-1):

- Some of these factors are in common with single user object interaction, for example immersion to create a feeling of presence or user interface for natural object manipulation
- Others are more related to collaboration, like the requirements that SHC place on intuitiveness of medium and interface (may in light of limitations in both), placing a need on the simulation to compensate. For example, adding a visual signal to substitute for the feeling of touch when an object is selected (see Virtual Gazebo in Chapter 5&6) or bending simulation time through space to hide the network delay [Sharkey et al., 1998].

A further outcome of this study is that many issues should be addressed to improve performance, handling and workflow of distributed collaboration. Therefore, the following section will discuss some future research directions that research community could take, to reach the goal of using collaborative VR in a wider range of activities.

9.1 Directions for Future Research

The research of this thesis focused mainly on the support for distributed closely-coupled collaboration using immersive display technology. While it showed that IPTs are useful for such collaboration, the research also raised some questions which could be answered in future research. For example, a major issue was the intuitiveness of the interface which resulted in issues while grasping objects and moving about within the immersive display. An alternative wireless interface with more precise tracking and a support for automatic constraint recognition could improve the ease of natural object interaction and lower the awareness of the technology. Some research in this area has been done and showed promising results [Marcelino et al., 2003; Osawa, 2006].

A further research direction might improve support for non-verbal communication through better gesture, posture and gaze support. The typical tracking interface allows only two to three sensors, to track for head and hand(s), which in combination with basic inverse kinematics give basic gesture support. Observations showed that participants made active use of this. However, more body sensors combined with enhanced avatar animation and support for facial expression could improve interactive communication and awareness of interactions from collaborators. Alternatively, a solution using video data of the actual user mapped onto the avatar could greatly enhance faithfulness and gain of trust.

Another problem of object interaction in immersive VR is adding the sense of touch through haptics interfaces. Although the research community for such devices is large, there are currently few haptic devices that would fit into a CAVE-like display. Furthermore, they can easily be seen which could detract from the illusion of presence.

A closer look into cultural differences between distributed groups, would further add to the understanding of how interfaces and application design have to be improved for global multicultural collaboration. In addition, most research is usually short term and a long term study would be beneficial to understand side effects of long term VR exposure as well as to identifying methods of adaptation (to technology, application and communication) used by participants.

Immersive CVEs appear promising for the support of distributed group collaboration and seem to be a good choice for studying people's behaviour during such collaboration, something demonstrated by this work. Applying the knowledge of the Virtual Gazebo research and related studies to create an application which could be used outside the academic research is desirable and Table 9-1 shows some potential application areas and their relevance to the usage of closely-coupled collaboration.

9.2 In final Summary

In an increasingly global economy there is increasing pressure to expand collaboration from co-located to geographically distributed groups, and this work has contributed to the understanding and further development of distributed closely-coupled collaboration, through immersive collaborative environments. This was approached primarily by measuring the impact of immersion and by isolating some key factors within this that have an impact on perceived and actual task performance. This work is contributing into the future development of distributed of collaboration by examining the strength and weaknesses of the technology supporting natural object focussed collaboration in distributed virtual environments and it is hoped that this work will generate a continued interest in many aspects of its content.

Description	Shared Objects				Use of objects			Location		Justification
	Solid	flexible	liquid	distance	handover	concurrent	sequence	fixed	moveable	
simulate emergency situations; train firemen	x	x		x	x	x	x	x	x	save lifes; reduces accidents; preparation for unexpected situations
training, planning, guiding of operations	x	x			x	x			x	save lifes; real training "objects" unavailable
conferences, emergency help, support with experts; (limited) training	x			x	x	x	x		x	involve distributed experts; real test not possible; only one try
planning of construction, archtecture; training/planning of construction tasks	x				x	x	x	x	x	Training; Demonstration; Test of new techniques
training, planning of dive "sessions" (sharing oxygen tank)	x				x	x	x		x	save lifes; better orientation; training
Demonstration for 3 people		x				x			x	demonstration; entertainment
Demonstration for 2 people - handing over a stick	x				x				x	demonstration; entertainment
training of staff for all kinds of situations; remote control in unreachable areas (small robot in collapsed tunnel to treat injured)	x	x			x	x	x		x	save lifes; reduces accidents; preparation for unexpected situations
make them accessible; allow different viewpoints; remote control	x					x	x	x		remote maintenance reduces costs
simulate physical/chemical processes; interaction; Research objects (eg. Molecules)	x	x	x		x	x	x		x	show invisible details; show/predict results
visualisation in real size; test ergonomy; collaborate/involve experts/managers	x				x	x	x	x	x	involve distributed experts; show final product early; save money
pilot may be out - no one to fly; training for emergency situations; remote control	x				x	x	x	x		remote control, help from ground
simulation of dangerous scenarios; training of soldiers; planning of strategies	x			x	x		x	x	x	simulate all thinkable events/situations
rehabilitation; remote doctor visits; remote treatment	x	x			x	x	x	x	x	better rehabilitation; reduce travelling costs
treatment by placing phobic patients into simulated situations	x	x	x	x	x	x	x	x	x	reduce travelling costs; visualization of objects causing phobias is save
simulate (unreal) scenarios/actions(move massive objects); collaboration of (not local) actors; interactive movies (place the audience into scene)	x	x		x	x	x	x	x	x	entertainment

Table 9-1: *Possible usage of distributed closely-coupled collaboration in various areas and a categorisation into the sharing and use of objects for these areas*

Appendix A – Questionnaires and Answers

The questionnaires in all chapters were aimed at ascertaining the user's subjective perception of collaboration, both generally and for a specific task. Questions were based on those of Usoh and colleagues [2000]. Answers could be given on a Likert-type scale [Sitzman, 2003] of 1-7, where 1 represented agreement to a very small extent and 7 to a very large extent. All data was obtained anonymously and participants signed a consent form (see below):

Consent form – agreement for voluntary user trial

surname/family name..…………………………..….

first names ……………………………………….

☐ Male ☐ Female

Please read this form carefully. If you have any further questions, do ask – we are here to help you. You have the right to change your mind at any time, including after you have signed this form.

I have been told that the following user trial does not involve any known health risk apart from a possible motion sickness/disorientation while conducting the tests. It is advised that no vehicle is operated the first hour after the trial.

I do not have a heart disease or epilepsy.

I understand that all data taken are anonymous and no personal data will be recorded or used.

I understand that this test is voluntary and I can drop out at any time without giving a reason.

Signature:….……………………… Date ..……………………..……….

Name (PRINT) ………………………..………

Appendix A.1 - Questionnaire of IPT-DT1-DT2 trial (Chapter 5)

Virtual Gazebo Questionnaire

Introduction

Project: **#H2.**
Coordinator: **#H3.**

Welcome to the Virtual Gazebo Questionnaire. To complete the survey, please click on the "Next Section" button in the center of the screen.

Please Note: You may need to use the scroll bar on the right hand side of the screen to read all the information on a page

To complete the survey, you may move from page to page using either

-- the buttons at the bottom of each page, or

-- the menu bar on the left side of the screen,

When you have completed this questionnaire, please press the "Exit" button below to save this information.

Please Note: **DO NOT RELOAD** - This would submit your data, delete all fields and you would need to start again.

About You

1. What is your sex?
   (CHECK ONLY ONE ANSWER)
   1. ○ Male
   2. ○ Female

2. How old are you?

3. Occupational Status
(CHECK ONLY ONE ANSWER)
   1. ○ Undergraduate Student
   2. ○ Masters Student
   3. ○ PhD Student
   4. ○ Research Assistant/Fellow
   5. ○ Staff - systems, technical
   6. ○ Faculty
   7. ○ Administrative Staff
   8. ○ Other

4. You were connected from which location (e.g. Salford, Linz)?

5. Optionally provide comments and your email address in the box.

About the Session

6. What is your Session number?

7. When did your session start?
(CLICK ON THE FIELD TO SET THE CURRENT TIME)

8. When did your session end?
(CLICK ON THE FIELD TO SET THE CURRENT TIME)

9. Please tell us which avatars were used by which user:

	CAVE (human-like)	Desktop (lego-like)	Desktop (human-like)
9a. Yourself ?	○	○	○
9b. USER 2 ?	○	○	○
9c. USER 3 ?	○	○	○

In the next few questions we would like to ask you about how you and the other users contributed, collaborated and performed during the task.

Contribution

Contribution: This tells us how much of the task was carried out by you and how you perceived the contribution of the other users. If you think you -contributed much more to a task than others, give yourself a high score relative to others.

To what extent did you feel each person **contributed** to the task of ...

(CHECK FOR EACH USER AND SCENARIO)

10. ... carrying a beam:

	1 (less)	2	3	4	5	6	7 (high)	No response
Yourself	○	○	○	○	○	○	○	◉
#9b.	○	○	○	○	○	○	○	◉
#9c.	○	○	○	○	○	○	○	◉

11. ... fixing a beam:

	1 (less)	2	3	4	5	6	7 (high)	No response
Yourself	○	○	○	○	○	○	○	◉
#9b.	○	○	○	○	○	○	○	◉
#9c.	○	○	○	○	○	○	○	◉

Performance

Performance: This question examines how well you were able to perform the task and how much you and the other participants hindered each other. For example, if someone removed a tool, this may have prevented you from carrying on with your task. Alternatively, if your partner gave you hints during the task, this may have increased the performance of the two of you.

12. Please give your opinion of how well you and the other person **performed** the task of ...

	1 (not at all)	2	3	4	5	6	7 (very well)	No response
... carrying a beam together	○	○	○	○	○	○	○	◉
... fixing a beam together	○	○	○	○	○	○	○	◉

To what extent did you feel each person **hindered** the other people from carrying out his/her task of ...

(CHECK FOR EACH USER AND SCENARIO)

13. ... carrying a beam:

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
Yourself	○	○	○	○	○	○	○	◉
#9b.	○	○	○	○	○	○	○	◉
#9c.	○	○	○	○	○	○	○	◉

14. ... fixing a beam:

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
Yourself	○	○	○	○	○	○	○	◉
#9b.	○	○	○	○	○	○	○	◉
#9c.	○	○	○	○	○	○	○	◉

Collaboration

Collaboration: Here we would like to know how much the two of you directly worked together.

To what extent did the two of you **collaborate** while ...
(CHECK FOR EACH USER AND SCENARIO)

15. ... carrying a beam:

	1 (not at all)	2	3	4	5	6	7 (a lot)	No response
Yourself and #9b.	○	○	○	○	○	○	○	◉
Yourself and #9c.	○	○	○	○	○	○	○	◉

16. ... fixing a beam:

	1 (not at all)	2	3	4	5	6	7 (a lot)	No response
Yourself and #9b.	○	○	○	○	○	○	○	◉
Yourself and #9c.	○	○	○	○	○	○	○	◉

Interface & Network

Interface & Network: With the following questions we would like to know to what extent the use of the system affected your interaction.. Our goal is to find out how important either the network, or interfaces with the system are in completing the task. For example, if you used a desktop computer and this made it quite hard to interact with the other people, please give a low rating. The same applies to delays from the network: for example, if it took some time to make an action visible, if you encountered problems picking up an object or if an action was not visible at the other end, and you experienced some inconsistencies.

17. To what extent did **the interface hamper** the task of ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... carrying a beam	○	○	○	○	○	○	○	◉
... fixing a beam	○	○	○	○	○	○	○	◉

18. How much have network induced **delays hamper** the task of ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... carrying a beam	○	○	○	○	○	○	○	◉
... fixing a beam	○	○	○	○	○	○	○	◉

19. How much have network induced **inconsistencies hamper** the task of ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... carrying a beam	○	○	○	○	○	○	○	◉
... fixing a beam	○	○	○	○	○	○	○	◉

20. Do you have comments which you like to metion in this correlation?

Social Interaction

Social Interaction: Working or talking with someone is already a form of social interaction. For example, normally when people collaborate on tasks, , some of their communication might be characterised as "small talk", but a good deal of it helps to coordinate team work. For us it is interesting how much you were able to use such social interaction while you performed the task. Please compare this with tasks performed in everyday life.

23. Please give your opinion of your ability to **socially interact** while ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... carrying a beam	○	○	○	○	○	○	○	◉
... fixing a beam	○	○	○	○	○	○	○	◉

Realism

Realism: The next questions are also related to the extent to which you were immersed inside the environment. We would like to know how realistic it was compared with performing a similar task in your everyday live. Did you feel like you were working with someone else and do you think this could be something like a real working place?

29. Did the interaction **appear realistic** to you?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
I think:	○	○	○	○	○	○	○	◉

30. Did you feel that you were **sharing the task** with another human being?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
I feel this:	○	○	○	○	○	○	○	◉

31. Did you feel that both of you were in the **same physical space**?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
Yourself and #9b.	○	○	○	○	○	○	○	◉
Yourself and #9c.	○	○	○	○	○	○	○	◉

Social feeling

Social feeling: In this section we would like understand the role of your emotions while performed our task. For example, if you felt you would have performed better with better social communication (more talk, collaboration, etc.) then give a low rating for the next question. Emotions are also a form of social communication: we use them constantly to express how we feel. These emotions can be expressed by the tone of someone's voice or by using gestures of the hands, head or body.

32. Did a **lack** of social feeling make the task harder?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
I think:	○	○	○	○	○	○	○	◉

33. Could you **sense** the emotions of the other persons?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
to #9b.	○	○	○	○	○	○	○	◉
to #9c.	○	○	○	○	○	○	○	◉

34. Did difficulties in **sensing** emotions make the task harder?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
with #9b.	○	○	○	○	○	○	○	◉
with #9c.	○	○	○	○	○	○	○	◉

Forms of SHC

35. To what extent did each of the following forms of Social Human Communication (SHC) **contributed to task performance**?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
Verbal communication	○	○	○	○	○	○	○	◉
Non-verbal communication	○	○	○	○	○	○	○	◉
Shared objects	○	○	○	○	○	○	○	◉
Environment	○	○	○	○	○	○	○	◉

Gestures

Gestures: When you work with someone, you normally use your voice to coordinate your actions, but subconsciously you use also your body to articulate what you want. These gestures help to support the spoken word. In virtual reality this kind of communication is only partially supported most of the time, and in some cases can hinder performance on a task. Therefore, we would like to know how much, if at all, you were able to use gestures and how important you found them.

37. Doing the experiment, how often did you need to **repeat a gesture** to make your request clear to the other person?

1 (not at all)	2	3	4	5	6	7 (very often)	No response

	1 (not at all)	2	3	4	5	6	7 (very often)	No response
I repeated gestures:	○	○	○	○	○	○	○	◉

Comments

40. Please provide any further comments you have on the positive aspect(s) of this trial.

41. Please provide any further comments you have on the negative aspect(s) of this trial.

42. Have you answered all questions?

    (CHECK ONLY ONE ANSWER)

    1. ○ Yes
    2. ◉ No

We like to thank you for your contribution to our research and hope that you had a good time.

When you have completed this questionnaire, please press the "Submit" button below to save this information.

Submit Cancel

	Display system user 1	CAVE																						
	Display system user 2	Desktop (lego like)																						
	Display system user 3	Desktop (human like)																						
$section[1]	Task 1: carrying beam																				M	SD	M %	SD %
$q[1]	To what extent did each person contributed to the task (carrying beam).																							
$qq[1][1]	-Yourself	5	4	6	7	5	5	5	7	7	7	4	NA	6	7	7	4	7	5	4	5.667	1.24	81	17.7
$qq[1][2]	-User 2	3	4	5	6	4	5	6	1	6	6	NA	3	3	7	7	4	7	4	4	4.722	1.67	67.5	23.9
$qq[1][3]	-User 3	3	4	1	3	2	NA	NA	7	6	6	NA	5	3	1	1	NA	7	4	4	3.8	2.08	54.3	29.7
$q[2]	Please give your opinion of how well you and the other person together performed the task (carrying beam).	4	3	5	4	5	5	4	7	7	7	6	4	NA	3	6	3	7	5	5	5	1.41	71.4	20.2
$q[3]	To what extent did each user hinder the task (carrying beam)?																							
$qq[3][1]	-Yourself	3	4	2	3	5	2	6	NA	1	1	3	NA	2	4	3	2	2	5	2	2.941	1.43	42	20.5
$qq[3][2]	-User 2	4	4	5	4	5	2	2	4	2	2	4	3	6	4	3	2	2	5	2	3.421	1.3	48.9	18.6
$qq[3][3]	-User 3	4	4	1	6	1	NA	NA	1	2	2	4	7	6	NA	1	NA	2	4	2	3.133	2.03	44.8	29
$q[4]	To what extent did the two of you collaborate?																							
$qq[4][1]	-User 2	2	4	5	5	1	6	2	7	6	6	5	NA	2	5	6	4	7	5	5	4.611	1.79	65.9	25.5
$qq[4][2]	-User 3	2	4	1	3	1	4	NA	7	3	3	4	NA	2	2	4	NA	7	4	4	3.438	1.75	49.1	25
$q[5]	To what extent did the interface hamper the task (carrying beam)?	5	5	1	4	5	2	6	1	1	1	4	4	5	3	NA	3	NA	5	4	3.471	1.7	49.6	24.3
$q[6]	How much have network induced delays hamper the task (carrying beam)?	5	5	2	7	4	3	3	3	2	2	3	7	5	5	NA	2	NA	5	5	4	1.66	57.1	23.7
$q[7]	How much have network induced inconsistencies hamper the task (carrying beam)?	5	7	2	7	5	2	2	NA	3	3	3	7	5	5	NA	2	NA	4	4	4.125	1.82	58.9	26
$section[8]	Task 2: fixing beam																							
$q[8]	To what extent did each person contributed to the task (fixing beam).																							
$qq[8][1]	-Yourself	6	5	5	5	6	6	4	7	6	6	5	NA	6	7	7	5	NA	5	2	5.471	1.23	78.2	17.6
$qq[8][2]	-User 2	4	4	5	3	4	6	2	4	6	6	5	3	2	7	7	1	7	5	5	4.526	1.78	64.7	25.4
$qq[8][3]	-User 3	4	4	5	3	2	4	NA	7	6	6	NA	5	2	4	7	1	7	4	5	4.471	1.81	63.9	25.8
$q[9]	Please give your opinion of how well you and the other person together performed the task (fixing beam).	6	3	4	4	6	5	2	7	7	7	4	NA	4	3	6	5	7	5	5	5	1.53	71.4	21.9
$q[10]	To what extent did each user hinder the task (fixing beam)?																							
$qq[10][1]	-Yourself	3	4	2	3	2	2	4	NA	2	2	5	NA	2	2	2	2	2	5	3	2.765	1.09	39.5	15.6
$qq[10][2]	-User 2	3	4	5	4	2	2	4	4	2	2	4	5	6	2	2	1	2	4	2	3.158	1.38	45.1	19.8
$qq[10][3]	-User 3	3	4	5	5	2	2	NA	1	2	2	NA	5	6	1	2	1	2	4	2	2.882	1.62	41.2	23.1
$q[11]	To what extent did the two of you collaborate?																							
$qq[11][1]	-User 2	3	4	5	5	6	6	3	7	6	6	5	NA	2	4	6	4	7	4	5	4.889	1.41	69.8	20.1
$qq[11][2]	-User 3	3	4	5	3	6	3	NA	7	6	6	NA	NA	2	1	4	1	7	4	5	4.188	1.94	59.8	27.7
$q[12]	To what extent did the interface hamper the task (fixing beam)?	5	NA	3	2	5	2	7	3	1	1	5	3	4	3	2	2	3	5	NA	3.294	1.65	47.1	23.6
$q[13]	How much have network induced delays hamper the task (fixing beam)?	5	NA	2	5	5	2	3	3	3	3	4	3	6	3	1	1	3	5	NA	3.353	1.46	47.9	20.8
$q[14]	How much have network induced inconsistencies hamper the task (fixing beam)?	5	NA	2	5	4	2	3	3	3	3	4	4	6	2	1	1	3	5	NA	3.294	1.45	47.1	20.7
$section[22]	general questions about collaboration with others																							
$q[22]	How social did it feel compared to the real world?	2	1	4	4	2	3	2	6	6	6	5	2	3	2	6	3	5	4	2	3.579	1.68	51.1	24
$q[23]	Did a lack of social feeling make the task harder?	2	6	2	3	3	5	5	7	2	2	3	2	6	7	3	2	4	4	6	3.895	1.82	55.6	26
$q[24]	Could you sense the emotions of the other people?																							
$qq[24][1]	-User 2	1	1	5	6	1	1	1	1	2	2	1	1	1	2	4	4	3	4	4	2.368	1.64	33.8	23.4
$qq[24][2]	-User 3	1	1	5	6	1	1	NA	2	2	2	1	1	1	1	4	1	3	4	4	2.278	1.64	32.5	23.4
$q[25]	Did difficulties in sensing emotions make the task harder?																							
$qq[25][1]	-User 2	2	3	2	3	2	5	6	5	2	2	4	2	5	1	2	2	5	4	5	3.263	1.52	46.6	21.7
$qq[25][2]	-User 3	2	3	2	3	2	5	NA	5	2	2	4	2	5	1	2	1	5	4	5	3.056	1.47	43.7	21.1
$q[26]	Did the interaction appear realistic?	1	2	4	2	2	6	2	7	3	3	NA	6	4	2	6	4	4	4	3	3.611	1.72	51.6	24.6
$q[27]	Did you feel that you were sharing the task with another human being?	2	5	4	6	3	7	2	6	5	5	6	5	3	4	NA	4	6	4	5	4.556	1.42	65.1	20.3
$q[28]	Did you feel that both of you were in the same physical space?																							
$qq[28][1]	-User 2	4	6	3	6	6	3	3	6	5	5	6	NA	2	5	7	5	7	4	3	4.778	1.52	68.3	21.7
$qq[28][2]	-User 3	4	6	3	6	6	3	NA	5	5	5	6	NA	2	2	7	1	7	4	NA	4.5	1.86	64.3	26.6
$q[29]	To what extent did each of the following contribute to task performance?																							
$qq[29][1]	Verbal communication	1	6	5	7	7	5	4	3	7	7	6	1	2	5	6	5	7	4	6	4.947	1.99	70.7	28.4
$qq[29][2]	Non-verbal communication	4	1	2	2	3	3	NA	7	7	7	5	1	2	1	2	2	NA	3	2	3.176	2.1	45.4	30
$qq[29][3]	Shared objects	1	5	4	5	3	5	NA	5	6	6	4	4	3	4	6	1	NA	3	5	4.118	1.54	58.8	21.9
$qq[29][4]	Environment	4	5	4	6	2	3	5	7	6	6	6	4	4	6	6	4	NA	4	5	4.833	1.29	69	18.5
$q[30]	Doing the experiment, how often did you need to repeat a gesture to make your request clear to the other person?	2	5	4	4	4	3	NA	1	5	5	3	NA	2	7	NA	2	NA	NA	3	3.571	1.6	51	22.9

Table A1-1: *Questionnaire results for CAVE-Desktop-Desktop user trial of chapter 5, results for CAVE user*

Display system user 1 — Desktop (lego like)
Display system user 2 — CAVE
Display system user 3 — Desktop (human like)

															M	SD	M %	SD %
$section[1]	Task 1: carrying beam																	
$q[1]	To what extent did each person contributed to the task (carrying beam).																	
$qq[1][1]	-Yourself	1	4	4	5	7	NA	2	4	7	7	5	6	3	4.58	1.98	65	28
$qq[1][2]	-User 2	7	4	6	4	7	7	5	7	7	7	3	7	5	5.85	1.46	84	21
$qq[1][3]	-User 3	5	2	4	4	1	NA	2	3	7	1	5	6	4	3.67	1.92	52	27
$q[2]	Please give your opinion of how well you and the other person together performed the task (carrying beam).	NA	3	3	5	5	2	4	5	7	6	5	7	4	4.67	1.56	67	22
$q[3]	To what extent did each user hinder the task (carrying beam)?																	
$qq[3][1]	-Yourself	1	3	5	2	2	NA	4	5	7	3	4	1	4	3.42	1.78	49	25
$qq[3][2]	-User 2	3	1	3	2	4	4	4	2	7	3	4	1	5	3.31	1.65	47	24
$qq[3][3]	-User 3	5	1	5	2	1	NA	4	NA	NA	1	4	1	5	2.9	1.85	41	26
$q[4]	To what extent did the two of you collaborate?																	
$qq[4][1]	-User 2	4	4	4	6	7	3	3	6	NA	4	5	7	3	4.67	1.5	67	21
$qq[4][2]	-User 3	3	4	5	6	1	NA	3	2	7	4	5	7	4	4.25	1.86	61	27
$q[5]	To what extent did the interface hamper the task (carrying beam)?	6	7	6	5	4	NA	4	7	5	4	3	3	4	4.83	1.4	69	20
$q[6]	How much have network induced delays hamper the task (carrying beam)?	3	7	7	4	2	7	4	6	5	5	3	2	4	4.54	1.81	65	26
$q[7]	How much have network induced inconsistencies hamper the task (carrying beam)?	5	7	6	5	3	6	NA	5	5	NA	3	2	5	4.73	1.49	68	21
$section[8]	Task 2: fixing beam																	
$q[8]	To what extent did each person contributed to the task (fixing beam).																	
$qq[8][1]	-Yourself	7	5	4	6	5	NA	2	5	7	7	3	7	4	5.17	1.7	74	24
$qq[8][2]	-User 2	5	4	5	5	7	NA	5	5	7	7	3	7	5	5.42	1.31	77	19
$qq[8][3]	-User 3	2	5	4	3	5	NA	2	2	7	7	5	7	5	4.5	1.93	64	28
$q[9]	Please give your opinion of how well you and the other person together performed the task (fixing beam).	6	7	5	4	5	NA	4	5	7	5	4	7	5	5.33	1.15	76	16
$q[10]	To what extent did each user hinder the task (fixing beam)?																	
$qq[10][1]	-Yourself	4	3	4	4	2	NA	5	3	7	4	3	2	4	3.75	1.36	54	19
$qq[10][2]	-User 2	5	3	5	4	5	NA	5	3	7	4	3	2	5	4.25	1.36	61	19
$qq[10][3]	-User 3	2	3	4	5	1	NA	5	3	7	4	3	2	5	3.67	1.67	52	24
$q[11]	To what extent did the two of you collaborate?																	
$qq[11][1]	-User 2	6	6	5	5	7	NA	5	5	NA	4	5	7	5	5.45	0.93	78	13
$qq[11][2]	-User 3	6	6	5	4	1	NA	5	5	7	4	5	7	5	5	1.6	71	23
$q[12]	To what extent did the interface hamper the task (fixing beam)?	7	7	6	6	5	NA	4	6	5	4	4	4	5	5.25	1.14	75	16
$q[13]	How much have network induced delays hamper the task (fixing beam)?	2	7	6	4	6	5	4	6	5	5	3	2	5	4.62	1.56	66	22
$q[14]	How much have network induced inconsistencies hamper the task (fixing beam)?	2	7	5	4	7	NA	4	6	5	NA	3	2	5	4.55	1.75	65	25
$section[22]	general questions about collaboration with others																	
$q[22]	How social did it feel compared to the real world?	3	1	3	4	2	5	5	2	6	3	3	5	3	3.46	1.45	49	21
$q[23]	Did a lack of social feeling make the task harder?	6	5	6	6	5	6	5	6	6	3	4	6	5	5.31	0.95	76	14
$q[24]	Could you sense the emotions of the other people?																	
$qq[24][1]	-User 2	4	2	3	6	1	2	5	1	7	1	2	6	3	3.31	2.1	47	30
$qq[24][2]	-User 3	6	2	3	2	1	2	5	1	7	1	2	6	3	3.15	2.12	45	30
$q[25]	Did difficulties in sensing emotions make the task harder?																	
$qq[25][1]	-User 2	6	6	5	3	3	2	4	6	2	NA	5	2	5	4.08	1.62	58	23
$qq[25][2]	-User 3	5	6	5	5	3	NA	4	6	2	NA	5	2	5	4.36	1.43	62	20
$q[26]	Did the interaction appear realistic?	2	2	3	5	5	2	4	2	7	3	3	4	3	3.46	1.51	49	22
$q[27]	Did you feel that you were sharing the task with another human being?	5	2	3	6	6	NA	4	6	7	7	4	6	4	5	1.6	71	23
$q[28]	Did you feel that both of you were in the same physical space?																	
$qq[28][1]	-User 2	5	4	4	6	5	NA	4	2	6	7	4	7	4	4.83	1.47	69	21
$qq[28][2]	-User 3	6	4	4	6	5	NA	4	2	6	7	4	7	4	4.92	1.51	70	22
$q[29]	To what extent did each of the following contribute to task performance?																	
$qq[29][1]	Verbal communication	4	7	7	4	3	NA	5	6	7	6	5	7	3	5.33	1.56	76	22
$qq[29][2]	Non-verbal communication	2	3	4	4	4	NA	5	1	NA	4	2	NA	3	3.2	1.23	46	18
$qq[29][3]	Shared objects	6	2	4	5	1	NA	5	4	NA	1	5	NA	4	3.7	1.77	53	25
$qq[29][4]	Environment	6	4	4	5	4	NA	NA	4	NA	5	5	NA	5	4.67	0.71	67	10
$q[30]	Doing the experiment, how often did you need to repeat a gesture to make your request clear to the other person?	6	4	3	4	1	NA	3	2	NA	NA	3	NA	5	3.44	1.51	49	22

Table A1-2: *Questionnaire results for CAVE-Desktop-Desktop user trial of chapter 5, results for Desktop user one*

Display system user 1 — Desktop (human like)
Display system user 2 — CAVE
Display system user 3 — Desktop (lego like)

														M	SD	M %	SD %
$section[1]	Task 1: carrying beam																
$q[1]	To what extent did each person contributed to the task (carrying beam).																
$qq[1][1]	-Yourself	3	5	1	1	NA	6	3	1	NA	4	5	7	3.6	2.2	51	31
$qq[1][2]	-User 2	3	3	6	7	7	NA	7	3	7	5	5	7	5.5	1.8	78	25
$qq[1][3]	-User 3	3	5	5	5	7	2	3	4	6	5	3	7	4.6	1.6	65	23
$q[2]	Please give your opinion of how well you and the other person together performed the task (carrying beam).	4	5	NA	5	5	6	1	1	NA	5	6	6	4.4	1.9	63	27
$q[3]	To what extent did each user hinder the task (carrying beam)?																
$qq[3][1]	-Yourself	4	1	1	3	2	1	4	1	6	5	3	3	2.8	1.7	40	24
$qq[3][2]	-User 2	4	1	2	5	NA	1	0	3	6	5	4	3	3.1	1.9	44	27
$qq[3][3]	-User 3	4	1	4	5	2	1	4	3	NA	5	3	NA	3.2	1.5	46	21
$q[4]	To what extent did the two of you collaborate?																
$qq[4][1]	-User 2	4	3	1	6	NA	5	NA	4	NA	4	5	6	4.2	1.6	60	22
$qq[4][2]	-User 3	4	5	2	6	6	1	3	4	6	4	4	NA	4.1	1.6	58	23
$q[5]	To what extent did the interface hamper the task (carrying beam)?	5	5	3	6	5	2	6	1	NA	5	3	5	4.2	1.7	60	24
$q[6]	How much have network induced delays hamper the task (carrying beam)?	6	3	1	5	6	2	3	1	NA	4	3	4	3.5	1.8	49	25
$q[7]	How much have network induced inconsistencies hamper the task (carrying beam)?	6	4	1	5	6	2	1	1	NA	4	3	5	3.5	2	49	28
$section[8]	Task 2: fixing beam																
$q[8]	To what extent did each person contributed to the task (fixing beam).																
$qq[8][1]	-Yourself	4	2	2	7	6	5	4	1	6	4	5	6	4.3	1.9	62	27
$qq[8][2]	-User 2	4	5	6	5	2	5	4	3	6	5	5	NA	4.5	1.2	65	17
$qq[8][3]	-User 3	4	5	6	5	6	4	4	4	6	4	3	6	4.8	1.1	68	15
$q[9]	Please give your opinion of how well you and the other person together performed the task (fixing beam).	4	4	5	6	6	5	3	1	7	5	6	6	4.8	1.6	69	23
$q[10]	To what extent did each user hinder the task (fixing beam)?																
$qq[10][1]	-Yourself	4	5	1	2	2	1	4	1	NA	5	3	4	2.9	1.6	42	23
$qq[10][2]	-User 2	4	4	1	4	2	1	1	3	6	4	3	4	3.1	1.6	44	22
$qq[10][3]	-User 3	4	3	1	4	2	1	4	2	NA	5	3	4	3	1.3	43	19
$q[11]	To what extent did the two of you collaborate?																
$qq[11][1]	-User 2	4	3	1	6	6	5	NA	3	6	5	5	6	4.5	1.6	65	23
$qq[11][2]	-User 3	4	3	1	2	3	4	5	3	NA	5	5	6	3.7	1.5	53	21
$q[12]	To what extent did the interface hamper the task (fixing beam)?	5	3	3	5	4	2	NA	1	NA	5	4	5	3.7	1.4	53	20
$q[13]	How much have network induced delays hamper the task (fixing beam)?	6	3	1	4	5	4	NA	1	NA	5	4	5	3.8	1.7	54	24
$q[14]	How much have network induced inconsistencies hamper the task (fixing beam)?	6	3	1	4	5	6	NA	1	NA	5	4	5	4	1.8	57	26
$section[22]	general questions about collaboration with others																
$q[22]	How social did it feel compared to the real world?	2	4	3	3	4	2	5	4	6	4	3	4	3.7	1.2	52	16
$q[23]	Did a lack of social feeling make the task harder?	6	6	7	5	5	5	7	2	4	3	4	7	5.1	1.6	73	23
$q[24]	Could you sense the emotions of the other people?																
$qq[24][1]	-User 2	2	1	4	2	1	2	5	3	6	4	2	5	3.1	1.7	44	24
$qq[24][2]	-User 3	2	1	2	2	1	4	4	3	3	4	2	5	2.8	1.3	39	18
$q[25]	Did difficulties in sensing emotions make the task harder?																
$qq[25][1]	-User 2	2	2	6	2	5	2	3	2	1	4	4	3	3	1.5	43	21
$qq[25][2]	-User 3	2	2	7	2	5	2	3	2	1	4	4	3	3.1	1.7	44	24
$q[26]	Did the interaction appear realistic?	3	2	2	5	5	1	5	3	7	3	4	5	3.8	1.7	54	24
$q[27]	Did you feel that you were sharing the task with another human being?	NA	2	3	6	4	6	3	4	7	3	5	6	4.5	1.6	64	23
$q[28]	Did you feel that both of you were in the same physical space?																
$qq[28][1]	-User 2	3	4	2	6	6	3	6	4	4	3	3	6	4.2	1.5	60	21
$qq[28][2]	-User 3	3	4	3	6	3	5	5	4	1	3	3	6	3.8	1.5	55	21
$q[29]	To what extent did each of the following contribute to task performance?																
$qq[29][1]	Verbal communication	5	5	5	1	5	5	4	3	6	3	4	6	4.3	1.4	62	21
$qq[29][2]	Non-verbal communication	3	2	2	3	3	3	4	2	2	3	2	NA	2.6	0.7	38	9.6
$qq[29][3]	Shared objects	4	2	4	2	NA	6	4	2	3	3	5	NA	3.5	1.4	50	19
$qq[29][4]	Environment	5	2	1	4	3	6	6	2	7	4	5	6	4.3	1.9	61	27
$q[30]	Doing the experiment, how often did you need to repeat a gesture to make your request clear to the other person?	4	2	NA	3	2	NA	4	1	3	4	2	NA	2.8	1.1	40	16

Table A1-3: *Questionnaire results for CAVE-Desktop-Desktop user trial of chapter 5, results for Desktop user two*

	Display system user 1	CAVE													
	Display system user 2	CAVE													
	Display system user 3	Desktop (lego like)													
$section[1]	Task 1: carrying beam											M	SD	M %	SD %
$q[1]	To what extent did each person contributed to the task (carrying beam).														
$qq[1][1]	-Yourself	3	7	2	4	5	NA	7	3	NA	7	4.75	2.05	67.9	29
$qq[1][2]	-User 2	6	7	NA	4	5	6	7	3	NA	6	5.5	1.41	78.6	20
$qq[1][3]	-User 3	3	NA	NA	NA	2	1	2	2	NA	3	2.17	0.75	31	11
$q[2]	Please give your opinion of how well you and the other person together performed the task (carrying beam).	5	5	4	3	6	6	6	NA	NA	6	5.13	1.13	73.2	16
$q[3]	To what extent did each user hinder the task (carrying beam)?														
$qq[3][1]	-Yourself	3	3	3	6	4	1	6	4	NA	2	3.56	1.67	50.8	24
$qq[3][2]	-User 2	5	3	NA	1	5	1	6	4	NA	4	3.63	1.85	51.8	26
$qq[3][3]	-User 3	3	NA	NA	NA	3	6	2	2	NA	6	3.67	1.86	52.4	27
$q[4]	To what extent did the two of you collaborate?														
$qq[4][1]	-User 2	4	7	NA	6	7	6	7	2	NA	6	5.63	1.77	80.4	25
$qq[4][2]	-User 3	4	NA	NA	NA	1	2	2	2	NA	3	2.33	1.03	33.3	15
$q[5]	To what extent did the interface hamper the task (carrying beam)?	4	5	NA	6	3	2	4	NA	NA	7	4.43	1.72	63.3	25
$q[6]	How much have network induced delays hamper the task (carrying beam)?	6	5	NA	5	2	2	4	NA	NA	5	4.14	1.57	59.2	22
$q[7]	How much have network induced inconsistencies hamper the task (carrying beam)?	2	5	NA	3	2	2	4	NA	NA	NA	3	1.26	42.9	18
$section[8]	Task 2: fixing beam														
$q[8]	To what extent did each person contributed to the task (fixing beam).														
$qq[8][1]	-Yourself	4	NA	NA	6	6	6	NA	NA	4	6	5.33	1.03	76.2	15
$qq[8][2]	-User 2	5	NA	NA	3	6	2	NA	NA	1	6	3.83	2.14	54.8	31
$qq[8][3]	-User 3	3	NA	NA	NA	1	1	NA	NA	NA	4	2.25	1.5	32.1	21
$q[9]	Please give your opinion of how well you and the other person together performed the task (fixing beam).	4	NA	2	4	6	5	NA	1	4	6	4	1.77	57.1	25
$q[10]	To what extent did each user hinder the task (fixing beam)?														
$qq[10][1]	-Yourself	3	NA	4	6	7	2	NA	1	2	4	3.63	2.07	51.8	30
$qq[10][2]	-User 2	4	NA	NA	1	NA	2	NA	1	1	4	2.17	1.47	31	21
$qq[10][3]	-User 3	3	NA	NA	NA	7	2	NA	1	NA	6	3.8	2.59	54.3	37
$q[11]	To what extent did the two of you collaborate?														
$qq[11][1]	-User 2	4	NA	5	2	2	6	NA	1	2	5	3.38	1.85	48.2	26
$qq[11][2]	-User 3	5	NA	NA	NA	6	1	NA	1	NA	4	3.4	2.3	48.6	33
$q[12]	To what extent did the interface hamper the task (fixing beam)?	3	NA	NA	5	2	1	NA	1	6	4	3.14	1.95	44.9	28
$q[13]	How much have network induced delays hamper the task (fixing beam)?	4	NA	NA	4	2	2	NA	1	2	4	2.71	1.25	38.8	18
$q[14]	How much have network induced inconsistencies hamper the task (fixing beam)?	3	NA	NA	3	2	2	NA	1	3	4	2.57	0.98	36.7	14
$section[22]	general questions about collaboration with others														
$q[22]	How social did it feel compared to the real world?	4	6	2	2	1	4	4	1	2	4	3	1.63	42.9	23
$q[23]	Did a lack of social feeling make the task harder?	3	2	4	6	1	3	4	5	2	5	3.5	1.58	50	23
$q[24]	Could you sense the emotions of the other people?														
$qq[24][1]	-User 2	5	2	3	1	1	3	1	1	1	1	1.9	1.37	27.1	20
$qq[24][2]	-User 3	5	NA	NA	NA	1	2	1	1	NA	1	1.83	1.6	26.2	23
$q[25]	Did difficulties in sensing emotions make the task harder?														
$qq[25][1]	-User 2	3	2	4	4	6	3	2	1	2	1	2.8	1.55	40	22
$qq[25][2]	-User 3	3	NA	NA	NA	6	3	NA	1	NA	1	2.8	2.05	40	29
$q[26]	Did the interaction appear realistic?	4	5	6	2	5	5	5	3	4	4	4.3	1.16	61.4	17
$q[27]	Did you feel that you were sharing the task with another human being?	NA	7	4	2	5	6	NA	4	7	6	5.13	1.73	73.2	25
$q[28]	Did you feel that both of you were in the same physical space?														
$qq[28][1]	-User 2	4	5	3	6	7	4	7	3	6	6	5.1	1.52	72.9	22
$qq[28][2]	-User 3	4	NA	NA	NA	2	3	7	3	NA	4	3.83	1.72	54.8	25
$q[29]	To what extent did each of the following contribute to task performance?														
$qq[29][1]	Verbal communication	6	2	4	1	1	5	7	1	1	7	3.5	2.59	50	37
$qq[29][2]	Non-verbal communication	5	5	3	3	7	4	3	1	5	4	4	1.63	57.1	23
$qq[29][3]	Shared objects	7	7	2	7	4	6	3	2	6	5	4.9	2.02	70	29
$qq[29][4]	Environment	7	5	4	6	5	5	2	5	5	1	4.5	1.78	64.3	25
$q[30]	Doing the experiment, how often did you need to repeat a gesture to make your request clear to the other person?	5	NA	4	NA	6	3	3	NA	1	4	3.71	1.6	53.1	23

Table A1-4: *Questionnaire results for CAVE-CAVE-Desktop user trial of chapter 5*

Appendix A.2 - Questionnaire of CAVE-CAVE trial (Chapter 6)

Virtual Gazebo Questionnaire

Introduction

Project: **#H1.**
Coordinator: **#H2.**

Welcome to the Virtual Gazebo Questionnaire. To complete the survey, please click on the "Next Section" button in the center of the screen.

Please Note: You may need to use the scroll bar on the right hand side of the screen to read all the information on a page

To complete the survey, you may move from page to page using either

-- the buttons at the bottom of each page, or

-- the menu bar on the left side of the screen,

When you have completed this questionnaire, please press the "Exit" button below to save this information.

Please Note: **DO NOT RELOAD** - This would submit your data, delete all fields and you would need to start again.

About You

1. What is your sex?
   (CHECK ONLY ONE ANSWER)
   1. ◯ Male
   2. ◯ Female

2. How old are you?
   ☐

3. Occupational Status

(CHECK ONLY ONE ANSWER)

1. ◯ Undergraduate Student
2. ◯ Masters Student
3. ◯ PhD Student
4. ◯ Research Assistant/Fellow
5. ◯ Staff - systems, technical
6. ◯ Faculty
7. ◯ Administrative Staff
8. ◯ Other

4. Optionally provide comments and your email address in the box.

About the Session

5. What is your Session number?

6. When did your session start?

(CLICK ON THE FIELD TO SET THE CURRENT TIME)

7. When did your session end?

(CLICK ON THE FIELD TO SET THE CURRENT TIME)

8. You were connected from which location?

(CHECK ONLY ONE ANSWER)

1. ◯ Salford
2. ◯ Reading

In the next few questions we would like to ask you about how you and the other users contributed, collaborated and performed during the task.

Contribution

Contribution: This tells us how much of the task was carried out by you and how you perceived the contribution of the other users. If you think you contributed much more to a task than others, give yourself a high score relative to others.

To what extent did you feel each person **contributed** to the task of ...
(CHECK FOR EACH USER AND SCENARIO)

9. ... fixing an object:

	1 (less)	2	3	4	5	6	7 (high)	No response
Yourself	○	○	○	○	○	○	○	●
the other user	○	○	○	○	○	○	○	●

10. ... carrying a beam:

	1 (less)	2	3	4	5	6	7 (high)	No response
Yourself	○	○	○	○	○	○	○	●
the other user	○	○	○	○	○	○	○	●

Performance

Performance: This question examines how well you were able to perform the task and how much you and the other participants hindered each other. For example, if someone removed a tool, this may have prevented you from carrying on with your task. Alternatively, if your partner gave you hints during the task, this may have increased the performance of the two of you.

11. Please give your opinion of how well you and the other person **performed** the task of ...

	1 (not at all)	2	3	4	5	6	7 (very well)	No response
... fixing an object together	○	○	○	○	○	○	○	●
... carrying a beam together	○	○	○	○	○	○	○	●

To what extent did you feel each person **hindered** the other people from carrying out his/her task of ...
(CHECK FOR EACH USER AND SCENARIO)

12. ... fixing an object:

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
Yourself	○	○	○	○	○	○	○	◉
the other user	○	○	○	○	○	○	○	◉

13. ... carrying a beam:

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
Yourself	○	○	○	○	○	○	○	◉
the other user	○	○	○	○	○	○	○	◉

Collaboration

Collaboration: Here we would like to know how much the two of you directly worked together.

14. To what extent did the two of you **collaborate** while ...

(CHECK FOR EACH USER AND SCENARIO)

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... fixing an object	○	○	○	○	○	○	○	◉
... carrying a beam	○	○	○	○	○	○	○	◉

Interface & Network

Interface & Network: With the following questions we would like to know to what extent the use of the system affected your interaction.. Our goal is to find out how important either the network, or interfaces with the system are in completing the task. For example, if you used a desktop computer and this made it quite hard to interact with the other people, please give a low rating. The same applies to delays from the network: for example, if it took some time to make an action visible, if you encountered problems picking up an object or if an action was not visible at the other end, and you experienced some inconsistencies.

15. To what extent did **the interface hamper** the task of ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... fixing an object	○	○	○	○	○	○	○	◉
... carrying a beam	○	○	○	○	○	○	○	◉

1 (not at all)	2	3	4	5	6	7 (very much)	No response

16. How much have network induced **delays hamper** the task of ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... fixing an object	○	○	○	○	○	○	○	●
... carrying a beam	○	○	○	○	○	○	○	●

17. How much have network induced **inconsistencies hamper** the task of ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... fixing an object	○	○	○	○	○	○	○	●
... carrying a beam	○	○	○	○	○	○	○	●

18. Do you have comments which you like to metion in this correlation?

Audio & Haptics

19. Please give us your opinion of how of much you were **relying on the audio** connection to complete the task successfully.

	1 (not at all)	2	3	4	5	6	7 (a lot)	No response
I have used it:	○	○	○	○	○	○	○	●

20. How much did you miss the **sence of touch** during the task?

	1 (not at all)	2	3	4	5	6	7 (a lot)	No response
I did miss it:	○	○	○	○	○	○	○	●

21. Do you think your **field of view (view frustum)** provided by the display was helpful and important to the task?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
I think:	○	○	○	○	○	○	○	◉

Social Interaction

Social Interaction: Working or talking with someone is already a form of social interaction. For example, normally when people collaborate on tasks, , some of their communication might be characterised as "small talk", but a good deal of it helps to coordinate team work. For us it is interesting how much you were able to use such social interaction while you performed the task. Please compare this with tasks performed in everyday life.

22. Please give your opinion of your ability to **socially interact** while ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... fixing an object	○	○	○	○	○	○	○	◉
... carrying a beam	○	○	○	○	○	○	○	◉

23. To what extent was your experience in working with the other person on this task **like the similar experience in the real world**, with regard to your sense of doing something together? While ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... fixing an object	○	○	○	○	○	○	○	◉
... carrying a beam	○	○	○	○	○	○	○	◉

---

being there together

Being together: Virtual reality normally tries to create an environment in which the user can immerse him or herself and may reach a point in the future where it is possible to forgets about the real world surrounding oneself. The following questions are an attempt to measure this idea by noting your perception of presence within the environment.

24. To what extent, if at all, did you have a sense of **being there** with the other person? While ...

1 (not at all)	2	3	4	5	6	7 (very much)	No response

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... fixing an object	○	○	○	○	○	○	○	◉
... carrying a beam	○	○	○	○	○	○	○	◉

25. To what extent did you **forget about the other person**, and concentrated only on doing the task as if you were the only one involved? While ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... fixing an object	○	○	○	○	○	○	○	◉
... carrying a beam	○	○	○	○	○	○	○	◉

Realism

Realism: The next questions are also related to the extent to which you were immersed inside the environment. We would like to know how realistic it was compared with performing a similar task in your everyday live. Did you feel like you were working with someone else and do you think this could be something like a real working place?

26. Did the interaction **appear realistic** to you?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
I think:	○	○	○	○	○	○	○	◉

27. Did you feel that you were **sharing the task** with another human being?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
I feel this:	○	○	○	○	○	○	○	◉

28. Did you feel that **both of you were in the same physical space?**

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
I think:	○	○	○	○	○	○	○	◉

Social feeling

Social feeling: In this section we would like understand the role of your emotions while performed our task. For example, if you felt you would have performed better with better social communication (more talk, collaboration, etc.) then give a low rating for the next question. Emotions are also a form of social communication: we use them constantly to express how we feel. These emotions can be expressed by the tone of someone's voice or by using gestures of the hands, head or body.

29. Did a **lack** of social feeling make the task harder?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
I think:	○	○	○	○	○	○	○	●

30. Could you **sense** the emotions of the other persons?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
I think:	○	○	○	○	○	○	○	●

31. Did difficulties in **sensing** emotions make the task harder?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
I think:	○	○	○	○	○	○	○	●

Forms of SHC

32. To what extent did each of the following forms of Social Human Communication (SHC) **contributed to task performance**?

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
Verbal communication	○	○	○	○	○	○	○	●
Non-verbal communication	○	○	○	○	○	○	○	●
Shared objects	○	○	○	○	○	○	○	●
Environment	○	○	○	○	○	○	○	●

Gestures

Gestures: When you work with someone, you normally use your voice to coordinate your actions, but subconsciously you use also your body to articulate what you want. These gestures help to support the spoken word. In virtual reality this kind of communication is only partially supported most of the time, and in some cases can hinder performance on a task. Therefore, we would like to know how much, if at all, you were able to use gestures and how important you found them.

33. To what extent, if at all, could you **use gestures**?

	1 (not at all)	2	3	4	5	6	7 (very often)	No response
I used gestures:	○	○	○	○	○	○	○	◉

34. Doing the experiment, how often did you need to **repeat a gesture** to make your request clear to the other person?

	1 (not at all)	2	3	4	5	6	7 (very often)	No response
I repeated gestures:	○	○	○	○	○	○	○	◉

Teamwork

35. To what extent, if at all, do you think it was easier to **work as a team or alone**?

	1 (alone is better)	2	3	4	5	6	7 (as a team)	No response
I think:	○	○	○	○	○	○	○	◉

36. To what extent was it **more effective** to work in a team?

	1 (useless)	2	3	4	5	6	7 (very effective)	No response
Teamwork was:	○	○	○	○	○	○	○	◉

Comments

37. Please provide any further comments you have on the positive aspect(s) of this trial.

38. Please provide any further comments you have on the negative aspect(s) of this trial.

39. Have you answered all questions and finished with the questionnaire?

(CHECK ONLY ONE ANSWER)

1. ◯ Yes
2. ◉ No

We like to thank you for your contribution to our research and hope that you had a good time.

When you have completed this questionnaire, please press the "Submit" button below to save this information.

Print

Submit

Cancel

QPL
Virtual Gazebo Questionnaire

				all combined				only IPT1		only IPT2				fast session (33%)			average session (37%)			slow session (	
What is your Session number?		M	SD	M %	SD %			M %	SD %	M %	SD %			M %	SD %		M %	SD %		M %	SD %
When did your session start?																					
When did your session end?																					
You were connected from which location?														percent			percent			percent	
Salford	time S			oneself				IPT1					19	33.3		21	36.8		17	29.8	
Reading	time R			other user						IPT2		novice only									
				6.8										4.7			6.4			9.6	
To what extent did you feel each person **contributed** to the task of fixing an object:						fixing															
Yourself		5.2	1.4	74.1	19.4			74.6	17.2	74.0	19.9			74.4	17.4		80.3	18.3		65.5	20.9
the other user		5.4	1.4	76.7	19.6			73.2	17.8	77.3	20.0			82.0	16.1		78.9	19.5		66.7	21.4
To what extent did you feel each person **contributed** to the task of carrying a beam:						carring															
Yourself		5.2	1.5	74.4	21.2			75.0	16.6	74.3	22.0			75.9	17.1		76.9	21.9		68.8	24.6
the other user		5.6	1.3	80.0	18.5			76.2	16.0	80.8	19.0			83.5	15.7		81.6	19.2		73.2	20.8
Please give your opinion of how well you and the other person **performed** the task of ...						performance															
... fixing an object together		5.5	1.2	78.2	16.4		fixing	66.1	7.4	80.2	16.7	85.7		86.5	13.4		77.6	15.4		71.4	17.9
... carrying a beam together		5.7	1.1	80.8	15.5		carring	76.2	12.4	81.6	16.0	81.0		86.5	13.2		83.0	13.3		72.3	16.1
To what extent did you feel each person **hindered** the other people from carrying out his/her task of fixing an object:						fixing															
Yourself		2.7	1.8	37.9	25.9			28.6	18.7	39.4	26.7			39.8	27.5		34.0	23.7		38.4	29.4
the other user		1.9	1.6	27.8	22.8			30.2	20.8	27.4	23.3			26.3	25.4		23.6	22.4		32.8	21.3
To what extent did you feel each person **hindered** the other people from carrying out his/her task of carrying a beam:						carring															
Yourself		2.7	1.9	38.2	27.3			25.0	14.8	40.3	28.4			44.4	30.8		34.7	24.6		34.8	27.1
the other user		1.9	1.7	27.3	24.0			22.2	14.5	28.3	25.3			25.6	26.5		28.6	26.6		26.9	19.5
To what extent did the two of you **collaborate** while ...						collaborate															
... fixing an object together		5.9	0.9	84.5	13.2		fixing	71.4	18.7	86.6	11.0	85.7		85.0	9.8		86.4	14.6		82.1	13.3
... carrying a beam together		5.9	1.0	84.7	14.7		carring	77.8	16.1	86.0	14.3	88.9		85.7	14.3		86.4	13.9		83.2	16.2
To what extent did **the interface hamper** the task of ...						interface															
... fixing an object together		3.1	1.5	44.4	21.7		fixing	50.0	20.2	43.4	22.0			36.8	17.1		46.3	25.1		49.5	21.5
... carrying a beam together		3.0	1.5	43.4	20.9		carring	42.9	20.2	43.5	21.2			41.3	18.0		39.5	19.6		50.0	22.1
How much have network induced **delays hamper** the task of …						delays															
... fixing an object together		1.5	0.9	22.1	13.5			17.9	6.6	22.7	14.3			18.8	6.9		25.9	18.4		21.9	11.9
... carrying a beam together		1.6	0.9	22.2	12.9			17.5	6.3	23.0	13.6			18.8	8.5		24.5	14.4		24.1	15.4
How much have network induced **inconsistencies hamper** the task of ...						inconsistencies															
... fixing an object together		2.1	1.5	29.6	22.0			20.4	7.6	31.0	23.1			21.1	8.8		35.0	27.6		34.7	23.6
... carrying a beam together		1.7	1.1	24.2	16.0			19.6	7.4	24.9	17.0			20.3	14.1		23.6	14.1		30.5	20.1
Please give us your opinion of how of much you were **relying on the audio** connection to complete the task successfully.		5.2	1.5	73.6	20.9			71.4	27.7	74.0	19.7			63.2	22.0		80.3	16.0		73.9	21.6
How much did you miss the **sence of touch** during the task?		3.6	1.9	50.8	26.8			55.6	25.2	50.0	27.3			54.9	27.7		52.4	27.6		46.2	26.5
Do you think your **field of view (view frustum)** provided by the display was helpful and important to the task?		5.4	1.4	76.5	20.3			76.2	20.2	76.6	20.6			76.7	19.1		78.2	19.5		71.4	23.1
Please give your opinion of your ability to **socially interact** while ...						social interaction		74.9	19.1	69.7	21.3			66.5			75.9			64.5	
... fixing an object together		5.0	1.4	71.9	20.2			76.8	20.1	71.1	20.3			69.9	21.3		76.9	19.4		64.3	18.1
... carrying a beam together		4.8	1.5	69.0	21.6			73.0	18.1	68.3	22.2			63.2	20.9		74.8	21.1		64.7	19.0
To what extent was your experience in working with the other person on this task **like the similar experience in the real world**, with regard to your sense of doing s						collaborative realism		68.0		65.1				64.3			66.7			63.2	
... fixing an object together		4.7	1.6	66.7	23.4			66.1	24.1	66.9	23.5			69.9	18.5		65.3	26.6		62.5	25.5
... carrying a beam together		4.5	1.8	64.4	25.8			69.8	20.8	63.4	26.6			58.6	22.9		68.0	25.9		63.9	27.7
To what extent, if at all, did you have a sense of **being there** with the other person? While ...						presence		77.3	18.9	68.1	21.9			65.8			72.8			67.0	
... fixing an object together		4.9	1.5	69.5	21.0			76.8	20.1	68.3	21.1			67.7	18.9		72.1	21.9		65.2	22.7
... carrying a beam together		4.9	1.5	69.5	22.1			77.8	17.7	68.0	22.6			63.9	24.1		73.5	22.3		68.9	20.4
To what extent did you **forget about the other person**, and concentrated only on doing the task as if you were the only one involved? While ...																					
... fixing an object together		2.7	1.7	38.4	24.3			50.0	24.1	36.6	24.0			39.1	25.1		38.1	25.7		36.6	20.2
... carrying a beam together		2.4	1.4	33.7	20.7			38.1	22.6	32.9	20.5			35.3	20.2		28.6	18.1		37.0	21.5
Did the interaction **appear realistic** to you?		4.0	1.6	57.4	22.6	realistic appearance		63.5	25.9	56.3	22.1			61.7	20.7		58.5	23.0		49.6	22.7
Did you feel that you were **sharing the task** with another human being?		6.0	1.0	85.5	13.9	co-presence		87.3	13.3	85.1	14.1			84.2	14.3		87.1	13.5		83.2	14.5
Did you feel that **both of you were in the same physical space?**		5.2	1.4	74.6	20.6	sharing space		81.0	16.0	73.4	21.2			75.9	17.7		74.1	23.7		71.4	20.8
Did a **lack** of social feeling make the task harder?		2.7	1.3	38.4	18.2			47.6	14.3	36.7	18.4			36.1	17.8		36.7	18.4		43.8	19.8
Could you **sense** the emotions of the other persons?		3.0	1.6	42.9	22.7	sense emotions		55.6	18.1	40.6	22.8			32.3	20.2		41.5	20.7		52.9	23.0
Did difficulties in **sensing** emotions make the task harder?		2.3	1.4	33.5	19.6			42.9	22.6	31.8	18.7			28.6	14.3		32.1	21.2		38.7	21.3
To what extent did each of the following forms of Social Human Communication (SHC) **contributed to task performance**?																					
Verbal communication		5.9	1.2	85.0	17.3	verbal		84.1	20.8	85.1	16.8			78.9	21.5		88.4	11.6		85.7	18.2
Non-verbal communication		3.1	1.5	43.6	21.8	non-verbal		30.2	15.1	46.0	22.0			47.4	23.0		42.9	20.7		42.0	23.4
Shared objects		5.0	1.3	71.7	18.8	objetcs		66.7	22.6	72.6	18.2			74.4	16.7		72.1	17.6		65.5	22.0
Environment		4.4	1.4	62.8	20.2	environment		73.0	23.1	60.9	19.3			62.4	17.7		64.3	20.5		58.0	21.7
To what extent, if at all, could you **use gestures**?		2.7	1.5	39.2	21.3			37.5	22.8	39.4	21.3			41.4	22.5		36.1	19.5		42.0	22.3
Doing the experiment, how often did you need to **repeat a gesture** to make your request clear to the other person?		2.0	1.4	28.0	20.6			30.6	25.3	27.6	20.1			22.3	15.1		28.6	19.6		33.0	26.5
To what extent, if at all, do you think it was easier to **work as a team or alone**?		5.4	1.6	76.5	22.2			66.7	24.7	78.3	21.5			75.9	21.9		79.6	21.0		70.6	24.5
To what extent was it **more effective** to work in a team?		5.7	1.3	81.8	18.0			73.0	20.8	83.4	17.2			80.5	15.5		88.4	18.4		73.1	17.4

Table A2-1: *Questionnaire results for CAVE-CAVE user trial of chapter 6*

Appendix A.3 -
Questionnaire of Single User trial (Chapter 7)

Virtual Gazebo Questionnaire

Introduction

Project: **#H2.**
Coordinator: **#H3.**

Welcome to the Virtual Gazebo Questionnaire. To complete the survey, please click on the "Next Section" button in the center of the screen.

Please Note: You may need to use the scroll bar on the right hand side of the screen to read all the information on a page

To complete the survey, you may move from page to page using either

-- the buttons at the bottom of each page, or

-- the menu bar on the left side of the screen,

When you have completed this questionnaire, please press the "Exit" button below to save this information.

Please Note: **DO NOT RELOAD** - This would submit your data, delete all fields and you would need to start again.

About You

1. What is your sex?
   (CHECK ONLY ONE ANSWER)
   1. ○ Male
   2. ○ Female

2. How old are you?
   [ ]

3. Occupational Status

   (CHECK ONLY ONE ANSWER)

   1. ◯ Undergraduate Student
   2. ◯ Masters Student
   3. ◯ PhD Student
   4. ◯ Research Assistant/Fellow
   5. ◯ Staff - systems, technical
   6. ◯ Faculty
   7. ◯ Administrative Staff
   8. ◯ Other

4. Optionally provide comments and your email address in the box.

About the Session

5. What is your Session number?

6. When did your session start?

   (CLICK ON THE FIELD TO SET THE CURRENT TIME)

7. When did your session end?

   (CLICK ON THE FIELD TO SET THE CURRENT TIME)

8. Please tell us on which machine you performed the tests:

	I used ...	No response
8a. CAVE ?	◯	◯
8b. Workbench ?	◯	◯
8c. Desktop ?	◯	◯

9. Do you normally have a good 3D vision?

   (CHECK ONLY ONE ANSWER)

   1. ◯ Yes
   2. ◯ No

10. Did you have a good 3D vision while using the CAVE and/or workbench?

   (CHECK ONLY ONE ANSWER)

   1. ◯ Yes
   2. ◯ No

In the next few questions we would like to ask you about how you performed during the task.

Performance

Performance: This question examines how well you were able to perform the task.

11. Please give your opinion of how well you **performed** the task of ...

	1 (poor)	2	3	4	5	6	7 (very well)	No response
... carrying an object in the CAVE	◯	◯	◯	◯	◯	◯	◯	◉
... fixing an object in the CAVE	◯	◯	◯	◯	◯	◯	◯	◉
... carrying an object on the Workbench	◯	◯	◯	◯	◯	◯	◯	◉
... fixing an object on the Workbench	◯	◯	◯	◯	◯	◯	◯	◉
... carrying an object on the Desktop	◯	◯	◯	◯	◯	◯	◯	◉
... fixing an object on the Desktop	◯	◯	◯	◯	◯	◯	◯	◉

Interface & Network

Interface & Network: With the following questions we would like to know to what extent the use of the system affected your interaction. Our goal is to find out how important either the network, or interfaces with the system are in completing the task. For example, if you used a desktop computer and this made it quite hard to interact with the objects, please give a low rating. The same applies to delays from the network: for example, if it took some time to make an action visible, if you encountered problems picking up an object or if an action was not visible, and you experienced some inconsistencies.

12. To what extent did **the interface hamper** the task on the ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... CAVE	○	○	○	○	○	○	○	◉
... Workbench	○	○	○	○	○	○	○	◉
... Desktop	○	○	○	○	○	○	○	◉

13. Do you think your **field of view (view frustum)** provided by the display was helpful and important to the task? While using the ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... CAVE	○	○	○	○	○	○	○	◉
... Workbench	○	○	○	○	○	○	○	◉
... Desktop	○	○	○	○	○	○	○	◉

14. To what extent was **the navigation** supporting the task on the ... While using the ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... CAVE	○	○	○	○	○	○	○	◉
... Workbench	○	○	○	○	○	○	○	◉
... Desktop	○	○	○	○	○	○	○	◉

Interface & Network

15. How much have network induced **delays hamper** the task on the ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... CAVE	○	○	○	○	○	○	○	◉
... Workbench	○	○	○	○	○	○	○	◉
... Desktop	○	○	○	○	○	○	○	◉

16. How much have network induced **inconsistencies hamper** the task on the ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... CAVE	○	○	○	○	○	○	○	◉
... Workbench	○	○	○	○	○	○	○	◉
... Desktop	○	○	○	○	○	○	○	◉

17. How much did you miss the **sence of touch** during the task on the ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... CAVE	○	○	○	○	○	○	○	◉
... Workbench	○	○	○	○	○	○	○	◉
... Desktop	○	○	○	○	○	○	○	◉

18. Do you have comments which you like to metion in this correlation?

being there

Being there: Virtual reality normally tries to create an environment in which the user can immerse him or herself and may reach a point in the future where it is possible to forgets about the real world surrounding oneself. The following questions are an attempt to measure this idea by noting your perception of presence within the environment.

To what extent, if at all, did you have a sense of **being there**? While using the ...

	1 (not at all)	2	3	4	5	6	7 (a lot)	No response
... CAVE	○	○	○	○	○	○	○	◉
... Workbench	○	○	○	○	○	○	○	◉
... Desktop	○	○	○	○	○	○	○	◉

Realism

Realism: The next questions are also related to the extent to which you were immersed inside the environment. We would like to know how realistic it was compared with performing a similar task in your everyday live. Did you feel like you were working in real environment and do you think this could be something like a real working place?

19. Did the interaction **appear realistic** to you? While using the ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... CAVE	○	○	○	○	○	○	○	◉
... Workbench	○	○	○	○	○	○	○	◉
... Desktop	○	○	○	○	○	○	○	◉

20. Did you feel that you were in the **a physical space**? While using the ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... CAVE	○	○	○	○	○	○	○	◉
... Workbench	○	○	○	○	○	○	○	◉
... Desktop	○	○	○	○	○	○	○	◉

21. Did a **lack** of social feeling make the task harder? While using the ...

	1 (not at all)	2	3	4	5	6	7 (very much)	No response
... CAVE	○	○	○	○	○	○	○	◉
... Workbench	○	○	○	○	○	○	○	◉
... Desktop	○	○	○	○	○	○	○	◉

Teamwork

22. To what extent, if at all, do you think it would have been easier to **work as a team or alone** inside this environment? While using the ...

	1 (alone is better)	2	3	4	5	6	7 (as a team)	No response
... CAVE	○	○	○	○	○	○	○	◉
... Workbench	○	○	○	○	○	○	○	◉
... Desktop	○	○	○	○	○	○	○	◉

23. To what extent was it **more effective** inside this environment to work alone? While using the ...

	1 (useless)	2	3	4	5	6	7 (very effective)	No response
... CAVE	○	○	○	○	○	○	○	◉
... Workbench	○	○	○	○	○	○	○	◉
... Desktop	○	○	○	○	○	○	○	◉

Comments

24. Please provide any further comments you have on the positive aspect(s)

this trial.

25. Please provide any further comments you have on the negative aspect(s) of this trial.

26. Have you answered all questions and finished with the questionnaire?

(CHECK ONLY ONE ANSWER)

1. ○ Yes
2. ◉ No

We like to thank you for your contribution to our research and hope that you had a good time.

When you have completed this questionnaire, please press the "Submit" button below to save this information.

Print

Submit

Cancel

QPL
Virtual Gazebo Questionnaire

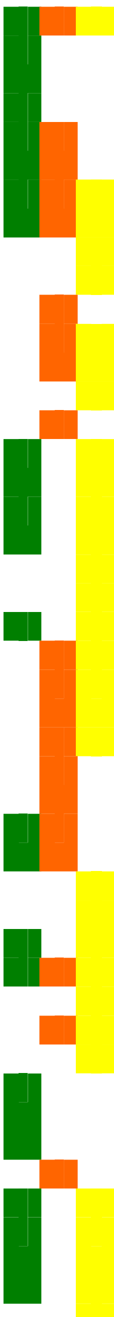
		1	2	3	4	5	6	7	8	9	10	11	M	SD	M %	SD %
What is your Session number?																
When did your session start?																
When did your session end?																
on which machine you performed the tests:																
CAVE	time C	7	6	6	5	9	6	9	8	12	6	6			7.3	2.1
Workbench	time W	6	5	6	8	7	5	10	10	7	5	8			7.0	1.8
Desktop	time D	7	9	6	5	6	5	7	4	7	5	6			6.1	1.4
Do you normally have a good 3D vision?		1	1	1	1	1	1	1	1	1	1	1	1.0	0.0	100.0	0.0
good 3D vision while using the CAVE and/or workbench?		1	1	1	1	1	1	1	1	1	1	1	1.0	0.0	100.0	0.0
Please give your opinion of how well you **performed** the task of ...																
... carrying an object in the CAVE		6	4	5	7	5	7	6	6	NA	6	5	5.8	1.0	81.4	13.6
... fixing an object in the CAVE		6	5	6	7	7	7	6	7	NA	6	7	6.4	0.7	91.4	10.0
... carrying an object on the Workbench		6	4	5	6	5	6	6	4	5	6	5	5.2	0.8	75.3	11.2
... fixing an object on the Workbench		6	4	6	7	6	6	6	5	4	6	4	5.6	1.0	77.9	14.8
... carrying an object on the Desktop		4	6	4	5	4	4	6	5	4	6	5	4.7	0.9	68.8	12.5
... fixing an object on the Desktop		4	6	4	6	4	5	6	5	NA	6	6	5.0	0.9	74.3	13.1
To what extent did **the interface hamper** the task on the ...																
... CAVE		2	4	2	1	3	2	3	3	1	5	3	2.3	1.0	37.7	17.2
... Workbench		3	3	3	2	3	3	4	5	1	6	5	3.0	1.1	49.4	20.6
... Desktop		5	3	5	6	6	5	5	4	1	5	2	4.4	1.6	61.0	23.1
Do you think your **field of view** was important during the interaction? While using the ...																
... CAVE		2	6	7	7	3	7	6	5	7	5	5	5.6	1.9	77.9	24.2
... Workbench		3	5	5	4	4	6	5	4	6	6	5	4.7	1.0	68.8	14.0
... Desktop		5	3	3	2	6	2	4	3	5	7	1	3.7	1.4	53.2	26.4
To what extent was **the navigation** supporting the task on the ... While using the ...																
... CAVE		6	5	5	7	5	2	6	4	6	6	5	5.1	1.5	74.0	19.0
... Workbench		5	4	5	7	4	5	5	3	7	6	3	5.0	1.3	70.1	19.6
... Desktop		3	4	5	3	3	7	4	3	6	6	6	4.2	1.5	64.9	21.5
How much have network induced **delays hamper** the task on the ...																
... CAVE		1	1	2	1	2	NA	1	1	1	2	1	1.3	0.5	18.6	6.9
... Workbench		1	1	1	1	3	NA	1	1	1	2	1	1.3	0.7	18.6	9.6
... Desktop		1	2	1	1	5	NA	1	1	1	1	1	1.6	1.4	21.4	18.1
How much have network induced **inconsistencies hamper** the task on the ...																
... CAVE		1	1	2	2	2	NA	4	1	1	1	1	1.8	1.0	22.9	13.8
... Workbench		1	1	1	2	2	NA	5	1	1	1	1	1.8	1.4	22.9	18.1
... Desktop		1	1	1	2	4	NA	5	1	1	1	1	2.0	1.6	25.7	21.1
How much did you miss the **sence of touch** during the task on the ...																
... CAVE		6	7	4	1	4	4	5	3	7	3	2	4.6	1.9	59.7	27.7
... Workbench		5	5	3	1	5	4	6	4	3	2	4	4.0	1.5	54.5	21.0
... Desktop		2	1	2	1	5	4	6	1	2	2	6	2.7	1.9	41.6	28.2
comments																
To what extent, if at all, did you have a sense of **being there**? While using the ...																
... CAVE		6	6	7	7	6	7	6	5	4	5	7	6.0	1.0	85.7	14.3
... Workbench		5	4	5	4	5	6	4	4	5	4	5	4.7	0.7	66.2	9.6
... Desktop		2	2	2	1	1	1	1	1	5	3	3	1.8	1.3	28.6	18.1
Did the interaction **appear realistic** to you? While using the ...																
... CAVE		6	5	5	5	6	5	6	3	6	5	6	5.2	1.0	75.3	12.9
... Workbench		5	3	3	3	5	4	5	2	5	4	4	3.9	1.2	55.8	14.9
... Desktop		1	1	1	1	1	1	4	1	4	3	4	1.7	1.3	28.6	20.2
Did you feel that you were in the **a physical space**? While using the ...																
... CAVE		5	5	7	7	6	7	7	5	7	6	7	6.2	1.0	89.6	12.9
... Workbench		4	3	5	6	5	6	5	3	7	5	6	4.9	1.4	71.4	18.1
... Desktop		3	2	1	1	1	2	3	1	5	3	3	2.1	1.4	32.5	18.2
Did a **lack** of social feeling make the task harder? While using the ...																
... CAVE		1	2	1	1	2	1	6	3	6	2	1	2.6	2.1	33.8	27.3
... Workbench		1	3	1	1	2	1	6	2	3	2	1	2.2	1.6	29.9	21.6
... Desktop		1	4	1	1	2	1	5	1	3	2	1	2.1	1.5	28.6	20.2
To what extent, if at all, do you think it would have been easier to **work as a team or alone**? While using the ...																
... CAVE		7	4	3	4	7	NA	7	2	7	6	7	5.1	2.1	77.1	27.9
... Workbench		1	5	3	2	4	NA	7	4	7	5	3	4.1	2.2	58.6	28.1
... Desktop		1	6	2	1	1	NA	6	1	1	2	1	2.4	2.3	31.4	29.2
To what extent was it **more effective** to work alone? While using the ...																
... CAVE		2	6	7	5	4	NA	4	5	2	4	5	4.4	1.8	62.9	22.5
... Workbench		6	5	7	5	3	NA	4	4	2	5	6	4.5	1.6	67.1	21.3
... Desktop		7	4	7	4	1	NA	3	7	6	6	7	4.9	2.2	74.3	30.0

Table A3-1: *Questionnaire results for single user trial of chapter 7*

Appendix B – Transcripts for Chapter 6

Session 46: *duration 4min* - ***fast session***

Transcript B-1: *Fast session example*



P A D	Time	Robin has both hands tracked while "Jeff" only has his right hand tracked. Robin has more experience of CAVEs based Virtual Gazebo building than Jeff.
	42:17	Robin activates to logging process from within the application, “ok, now we can do a proper session” (referring to the previous training)
	42:21	(both are looking at each other)
	42:24	Jeff: “Ok, so it is getting the t-joint first” (he is standing in front of the t-joint material stack)
	42:28	Robin: (looking at Jeff) “yap, do you put it on one side and” (rotating around, towards the tools) “I get the drill and fix it”
	42:34	Jeff is picking up a t-joint a moves towards a beam
	42:37	Robin picks up the drill and moves back to Jeff
	42:39	Jeff “its”, Robin “yea like this, exactly” (seeing Jeff holding the joint at a beam)
	42:40	Jeff stepping forward and stretching his arm to position the t-joint, Robin moving to Jeff “okidoke”
	42:44	Robin is at Jeff’s location, “hols” and used the drill
	42:48	Jeff see’s Robins success “right”
	42:49	Robin moves back to the tools “and I get a screw”, picking up the screwbox and returning it to Jeff
	42:55	Robin (facing Jeff) creates a screw, pickes it up (after two trials) “put it in here” (pointing with his second hand)
	43:05	Robin “mneahh” (the screw dropped down), Jeff “uhh”
	43:11	Robin creates a new screw
	43:17	and inserts is successful “ok” (moving backwards) “and the screwdriver is here”
	43:22	Robin picks up the screwdriver and moves back to Jeff “almost done”
	43:25	Robin holding up the screwdriver and looking at Jeff “yea, are you ready”, Jeff “alright”
	43:28	Robin fixes the t-joint “ok, yeap, fixed”
	43:30	Robin moves back and see’s Jeff still holding on to the joint “ok”, Jeff lets go “aehha” <surprised>
	43:32	Jeff: “right”, Robin “cool, so I hold the other one then now” (moving towards the t-joint stack)
	43:38	Robin picks up a t-joint while Jeff takes the drill
	43:44	Robin is moving to the second beam, Jeff looks at him and follows his action
	43:48	Robin positions the joint “like this”, Jeff moves around Robin to drill
	43:53	Jeff drilling “upp”, Robin agreeing “yap”, Jeff “is that done?”, Robin “ja, that’s ok”
	43:59	Jeff puts drill aside and moved to the screwbox
	44:03	Jeff gets a screw while Robin looks to his left to watch Jeff
	44:12	Jeff has a screw and rotates 180° (facing now the other screen) and moves to Robin
	44:13	Robin is using his left hand (tracked) to point where the screw should go “you need to put it somewhere here now, because you had your holes here, down there” in the meantime Jeff inserts the screw
	44:22	Jeff “ok like that”, Robin “yep”, Jeff releasing the screw, Robin “okidoke”
	44:24	Jeff moving back to find the screwdriver “and screwdriver is there”
	44:28	Jeff stretching his arm forward to pick up the tool, but hitting the CAVE screen “move better forwards” and moves with the joystick, Robin is watching this over his shoulder
	44:37	Jeff is back with the screwdriver and rotates with his body left (avoiding Robins

avatar, to step beside him) to used it,

44:42 fixing successful, Robin “okee” (moving back), Jeff I thing that’s it”, Robin “yep, I’ll let go”

44:49 Jeff “there we go”, Robin moving toward the tools “so now we get a beam in between”

44:53 Jeff “right, so the tools are over here” (moving to the carry tools)

45:05 both have a carry tool, robin moves to the right side and Jeff (seeing Robin right) to the left

45:10 both bend down to pickup the beam and lift it

45:15 Jeff moving backwards “ups to far” seeing Robin closing in on one t-joint, Robin looks back to see Jeff <laughs>

45:21 inserting the beam, both looking at each others action

45:28 Robin finished “yeap okidoke”, Jeff inserted his end to low, Robin see that “yea you take it again and put it in the middle part” (pointing with his hand)

45:35 Jeff “that’s fine, get the” carry tool

45:40 Jeff has difficulties to pickup the tool, Robin wants to help with his tool “I can take this one here as well”, but Jeff is faster “there we go”

45:47 Beam is now positioned correctly, Robin “yea like this”, Jeff “yea, there we go”

45:51 Jeff is moving back to see the side, Robin “I get the drill” picking it up “and you can get a screw ready”

45:55 Jeff “right, where are, the screw are around here, aren’t they” (looking around), Robin see’s Jeff “yea, there” pointing in Jeffs direction with the screwbox near him, Jeff “are there they are”

46:00 Jeff picks up a screw and Robin moves to the screwdriver “get a screwdriver already”

46:09 Jeff at the beam with a screw “so it was here?”, Robin standing on the other side of the beam “yea it doesn’t matter really”

46:12 Jeff inserts the screw and Robin immediately fixes it

46:16 Robin “okidoke, that’s it, finished” Jeff “yea”

Session 20: *duration 7min* - ***average session***

Transcript B-2: *Average session example*

P A D Time Sara (non-native English speaker) in Salford has both hands tracked while Jim (native English) in Reading has only one hand. Both are novice users and had 15-30 min training.

48:56 Jim: <picks up a joint> “So a yeah, are we fixing this in first, the t-joint” no response from Sara as she gets still get some instructions how to pickup objects

49:05 Jim: “ok can you find the drill”

49:06 Sara: “ja”

49:15 Sara: “ok where are you” <looking around> with her head but also using the joystick Both user face a wall (right and left respectively)

49:17 Jim: <aligns the t-joint along one of the beams>

49:30 Sara: <drills a hole> and then <drops the drill>

49:49 Sara: moves back to the tool area and picks up the screw box and carries it to Jim and the beam

50:08 Sara: <creates screw> and tries to insert it <laughs> Jim “ok” <laughs> screw position is not perfect, she tries to realign

50:21 Sara: “I think it is not in the beam”

50:25 Sara: finds a position, Jim “yea yea pick that” Sara “ok, ok” Jim “that should do it”

50:31 Sara: takes a step backwards and looks at the alignment “I don’t know, but it looks ok”

50:34 Jim: “what are you up to? Screw it in first, yea”

50:36 Sara <rotates> 180 degree, now facing the right wall, then <turning back> looking where to find the screwdriver,

50:45 finally she spots it and <moves> toward it (with joystick), multiple tries to pickup the screwdriver,

51:11	picking up did not work, she moves further (now facing front screen)
51:16	Sara: “ah, no, I try to take it”
51:16	Jim: “so, I am waiting here, are you alright?”
51:19	Sara: “ja ok” <holds the screwdriver> and is looking where to move (with her head looking around)
51:23	Jim: “have you got the screwdriver?”
51:24	Sara: “yea”
51:25	Jim: “ok” Sara <rotates> with joystick and moves back to Jim
51:35	Close to Jim, Sara <starts to walk> the last meter and <inserts> the screwdriver
51:36	Sara: “ok, ok” after fixing (joint and screw with beam) “you can let go now” while <stepping back> in CAVE
51:40	Jim: “right”, here we go” after holding the t-joint all this time he is now <taking a step back> and drops his (holding) hand
51:42	Both look at the result, which looks straight
51:43	Jim: “are you going to hold up the other one this time?”
51:45	Sara: “ok”
51:49	Sara moves to the joint stack and picks up another t-joint
51:52	Jim looks for the drill and picks it up, then moves with joystick back to have a look at both beams
51:55	Sara is on Jims front screen, Jim <glances> at her
51:58	Sara: rotates with joystick so that the constructions site is on her left CAVE wall
51:59	Sara: is moving towards the second beam, Jim still looking at her point with his hand (which holds the drill) to the beam “I guess it comes right into the middle”
52:05	Sara <overshoots> the target, ends up at the first beam, is confused and looks around
52:09	In the mean time, Jim sees Sara near the beam and says “right get them about the same hight” and <points with the dill> “so there”
52:12	Sara: “aeh where is it” (beam is on the open side of the CAVE), rotates and finds it
52:21	Sara is moving in the CAVE to align the t-joint with the beam and Jim asks instructor “and then what do I press to use it again?” he just need to intersect the drill with the joint
52:28	and now Jim is <doing> it “got it right”
52:32	Jim: moving backwards “getting some; will see”
52:35	Sara can see the screwbox near the other beam “the whole box is on the other side” “can you see it, yea”
52:43	Jim “yea, yea got it” he creates a screw and picks it up
52:47	Jim moves with screw to Sara and tries to insert it, but drops screw “up, shouldn’t have to do that” and <looks down>
52:50	Jim <bends> down and <picks up> the screw,
52:54	Sara “perhaps is better to take the other one” she glances at him
52:57	Jim: “what?, in the there” insert successful “ah yea, here we go”
53:03	Jim moves to find the “ah screwdriver” and Sara follows his movements (looking with her head)
53:08	Jim <overshoots> the tool “upp” and physically rotates to get into a better pickup position
53:14	Jim picks up the screwdriver and moves back to Sara and the joint
53:21	Sara looks at him “na”
53:22	Jim fixes the structure “here we go”
53:25	Jim <moves> back “is that it”, Sara “ok” and she releases her hold on the joint
53:27	Jim: “yeap, alright we are going to the picking up tools” while moving towards them and picking one up
53:32	Sara follows and picks one up as well
53:40	Sara sees Jim to her right and chooses the left beam side and gets closer to the first beam, meanwhile Jim navigating to find the beam side, he sees her at one end and chooses the other
53:41	Jim glances at Sara, now both bend down
53:42	Sara: “ok” (she got one end), Jim dropped his tool “sorry”, picks up the tool and tries again “here we go”, meanwhile Sara stays bend down and waits for Jim
53:49	Both synchronously straighten and thereby lift the beam, Jim “got it”
53:51	Jim takes the lead and moves back to the construction side, Sara is looking where he is going and follows
54:05	Jim tries to insert his (moves the last meter in the CAVE), but drops it “ups” now he

	tries to pick up the beam directly, but instructor tells him to use the tool, meanwhile Sara inserts her end of the beam (but is now very close to the right wall and out of the video frame) and says "ok, its ok"
54:24	Jim picked up the beam again "right"
54:28	Jim inserts the beam, but a the lower end (instructor tells him to use the middle), Jim "ok I see what you mean", takes the beam and inserts correctly, Sara "ok fine" Jim "here we go"
54:35	Jim: "and then" he picks up the screwdriver
54:50	Sara sees Jim with a screwdriver and comes with a screw, Jim "and if you put that in"
55:00	Sara moves back, next to Jim, in front of them they see the structure and the drill, Sara "do you have the other one at first, I think" Jim "ok" <laughs>
55:04	Jim "sorry what"
55:04	Sara "I can do it" and moves to find the drill, bends down to go underneath the beam but <overshoots>
55:09	Meanwhile Jim says "ok, you couldn't couldn't get"
55:12	Sara is trying to reorient herself but gets frustrated "man yea, I lost the thing, sorry" Jim "ok" <smiles>
	(Sara does the following seen completely by moving physically around and not using the joystick)
55:22	Sara goes back to the beam, find the drill moves physically around the beam and picks up the drill
55:31	Jim is looking what she is doing "up" and now realises they have to drill a hole first before the can continue "ah alright, drill a hole first, I didn't realise"
55:37	Sara drills a hole "ok now we" and moves to get a screw,
55:45	<creates one>
55:50	<picks it up>
55:55	and <inserts it> "hm and you", meanwhile Jim gets the screwdriver "here we go" and fixes the whole structure
56:06	Sara: "fine" Jim "yeap, there we go"

Session 43: *duration 15min* – ***slow session***

Transcript B-3: *Slow session example*

P	A	D	Time	
				John (native English speaker) in Salford has both hands tracked while Shawn (native English)in Reading has only one hand. Both are novice users and had 15-30 min training.
			41:52	John is still busy with his setup, Shawn takes the drill and starts drilling on one beam
			41:58	Shawn: turns in direction of John and asks "What do you want me to do?" <pause> "Do you .." "Do you want to pick up the t-joint?"
			42:04	John: "alright" is picking it up "got it"
			42:10	Shawn sees John but also the building side (on the left wall) and says "on there" pointing in the direction a one of the beams, meanwhile John is looking at Shawn (right wall) (joint is on the ground)
			42:14	Shwan: "aeh you need to hold it up"
			42:17	John: "aeh, have I dropped it yea" Shawn is gazing at john and observes how he is trying to pick up the joint
			42:26	John picks up the joint and Shawn gets closer to the beam, pointing to a position on the beam "if you keep it there"
			42:42	John sees the gesture and holds the joint in position "like that?"
			42:44	Shawn: "and" start moving away towards the tool stack "yeap" (in response to Johns question), John "just like that" (John sees now only his joint and the beam)
			42:45	Shawn is now looking for the screwbox (while he moves he also rotates with his body as this offers him a better view)
			42:49	Shawn: "I need .." standing in front of the screwbox, creating one and picking the screw up (right wall)
			42:56	Shawn: rotates to the left wall as he can see there the John and the structure, then

moves towards him
43:05 Shawn inserts the screw (but he needs to drill first, which he has not realised yet), the screw does not stick “is that not”
43:15 (The instructor tells him that he needs to drill a hole first, meanwhile John is instructed how to hold the joint in a better position)
43:25 Shawn picks up the drill (which is in front of him) and drills a hole, then bends down and picks up the screw, he inserts the screw
43:28 (John holding the joint upright now)
43:41 Shawn <rotates> around to go back to the tools and pick up the screwdriver
43:43 John: “is it better now mate?”
43:47 Shawn “aehh, its steady yea, I can’t see the screw now” and he starts moving around (with joystick) “I just try from a different angle”
43:54 Shawn gets closer to John and moves his head to look around the beam
44:06 (In the confusion both drop there item)
44:07 Shawn: “dropped it”
44:10 John: “oh no that’s ..”
44:12 While John is picking up the joint, Shawn moves back to the screwbox and gets another screw “I’ll get another one”
44:22 John is holding up the joint and asks “do you wanna really now” Shawn looks at him (with a screw in his hand)
44:28 Shawn: “I can see it now” sees the screwdriver in front of him and picks it up, he tries to fix the structure “yep” <pause> “I think you can release it now”
44:40 John lets go but it is not fixed and the joint fells down
44:41 Shawn: “NO”
44:46 John looks down trying to get hold of the joint again
44:50 Shawn: “if, if you pick it up again” watching and gazing at John
45:02 John has difficulties in picking up, Shawn sees the problem and intervenes by picking up the joint himself
45:11 He is trying to show John were to position the joint “if you” but then drops it himself
45:15 John sees that the joint is dropped again and looks for it on the ground
45:17 Shawn: “its behind you”
45:21 John moves backward (with joystick) “alright”
45:25 John picks up the joint and Shawn tries to get the screw,
45:34 but realises that he needs to drill first “oh yea” <pause> “started over”
45:41 Shawn is drilling “uhh” <laugh> “that was a bit close” and he puts the drill aside
45:51 Shawn moves back looking for a screw, found one and thinks he picked it up and inserts it, but “I think I dropped the screw” “I did”
46:08 His third attempt inserting the screw does not work, because the joint is not correctly aligned (the instructor makes him aware of this) “the middle part of the T has to face towards the other post” at the same time he is gazing between John and the other beam (further to his right)
46:37 John: “ah, do you mean” rotating the joint and moving around
46:45 John: “alright I got you know” back at the beam trying to align
46:51 Shawn: seeing John and the joint “its gotta face towards me”
46:55 John: “where are you? I can’t see you” (sees only the beam and joint)
46:59 John: looks to his right side “alright”
46:59 Shawn “if you walk around the beam” (making a gesture with his arm indicating how to walk)
47:05 John runs off and Shawn view follows him (by moving his head) <laugh>
47:10 Shawn sees John coming back and points (the drill still in hand) at the beam
47:15 John sees Shawn, which is close to the second beam and indicates that John could use this one and the direction the t-joint should face (without saying a word)
47:20 John acknowledges by “le me drop it” (in order to get a better grip on the joint)
47:29 John has the joint again and Shawn indicates once more
47:43 John is now getting closer to the beam and aligns the joint “is that any better?”
47:45 Shawn already drilling a hole “that’s perfect”
47:47 Shawn then is busy finding one of the screws, inserts it (moving close to John who can observe him)
48:04 and next looks for the screwdriver
48:17 While fixing the joint Shawn gets very close to John “up” and moves immediately

back, but the joint is fixed “aeh, you can let go now”
48:23 John releases the joint (which stays in place) “right” and looks at it “cool” <laughs>
48:27 Meanwhile Shawn moves already to the material stack to get the second t-joint “I’ll take this time and you, aehh” “and you screw it in” having the joint already in his hand
48:35 John: “right where is the box” looking around (but his instructor tells him he needs to drill first) Shawn “blue one”
48:43 John: “which one is the drill?” <looking around>
48:45 Shawn is looking to find the drill (which is in front of him)
48:48 John sees Shawn coming towards him and a tool in front of himself, he points at it “the one?”
48:48 Shawn: “the drill which is just here” pointing at it and at the same time gazing between John and the drill
48:57 Shawn moves on to align his t-joint and John picks up the drill
49:01 John is left of Shawn “now you need” and starts drilling
49:06 John moves back to get a better angle, but then Shawn “that’s fine”
49:14 John keeps drilling and Shawn says “you need a screw now”
49:18 John: “right let me go and get a screw” moving away trying to find the screwbox
49:25 Shawn: studies his joint “don’t know if that’s level”
49:30 Shawn now looks to his right to see what John is doing, he sees that John is in front of the screwbox and says “it’s the blue one”
49:35 John: “yeah” while picking up the whole box and bringing it back to the construction side right behind Shawn
49:41 John is creating a screw and tries to pick one up, Shawn keeps looking behind him to see what John is doing but also looks to his alignment of the joint
John has problems to pickup a screw
50:22 John: “I am bringing a screw, oh dropped it”
50:25 Shawn: “ok”
50:39 John tries to insert something (and now gets advice from his instructor)
50:48 Shawn: looking around the beam “I think you dropped it”
50:51 John “yeah” and is now taking another screw and tries to insert it
51:17 Shawn: making a gesture with his non-tracked hand to have the screw further down “ahh, the screw is not touching the wood I think”
51:25 Meanwhile John is looking for the screwdriver
51:32 Shawn gets the information that “ok apparently it is fine”
51:35 John has difficulties to pick up the screwdriver and his instructor shows him how to do it, Shawn is unaware of that but follows Johns avatar (with his head), consequently he drops the joint <is annoyed> “ach” “is that ok or does it have to be done again?” (asking the instructor)
52:05 Shawn picks up the joint
52:09 John: “right got it”
52:12 Shawn: “just changing the way I hold it”
52:23 John is coming back but <overshoots> and rotates around
52:30 John is fixing the structure
52:34 Shawn: “done”
52:38 Structure is fixed, but John keeps using the screwdriver
52:42 Shawn: “aeh, we need a beam now” he moved already to the carry tools
52:45 now John is confused not sure what to do (gets further instructions from his instructor)
52:55 John gets closer to the carry tools, while Shawn is changing the grip on his tool
53:07 Shawn sees John near a tool “that’s the one” and moves away to find a beam
53:14 John is not far from a beam and uses his tool with on end of the first beam
53:15 Shawn sees the selection and uses his tool with the other end
53:18 John is still bend down
53:23 Shawn waves with his tool to indicate that the beam is now moving
53:24 John gets up
53:28 Shawn moves away to carry the beam over to the structure
53:32 John follows slowly
53:40 Shawn manages to insert his end of the beam in one of the t-joints and observes how John is doing the same at his end

53:50	for a moment both are not sure what to do next, John “right”, Shawn “do we need to nail it in?” (instruction follows to screw it)
53:59	John is near the screwbox “screw”
54:03	Shawn looks at John “you get the screw I’ll get the drill”
54:05	John: “right, ok”
54:11	While John it trying to get a screw, Shawn is locating the drill and brings it back to the structure (slight back moving reaction when he <overshoots> to close to the beam)
54:30	Shawn gets close to John and uses the drill
54:32	John sees Shawn “done it already” “have you done the holes?”
54:36	Shawn: “not yet” “there you go”
54:45	John: “anywhere?” inserting the screw, meanwhile Shawn tries to locate the screwdriver
54:56	John: “I just, .. got it?” Shawn “no”
55:01	Shawn is near the screwdriver but has difficulties picking it up
55:09	John moves back to better see what Shawn is doing and how the structure is looking
55:13	John: “right” <pause> “where is the screwdriver”
55:18	Meanwhile Shawn got hold of it already and is going to use it
55:23	John: “right, yeaa, he’s got it”
55:31	Shawn is trying different locations to fix it nothing happens “Did you put a screw in?”
55:32	John observing Shawn “yea, yea it is in the middle” and moves closer to point out which location he is talking about “inside”
55:36	Shawn uses the screwdriver and the beam is fixed
55:42	After switching off the logging (as advices by instructor)
55:50	Both wave with their hand
55:57	Shawn: “bye, bye”
55:58	John: “cu you later”<laughs>

References

Amirabdollahian, F. (2004). *An investigation of robot-mediated therapies and therapy effects on the recovery of upper limb post-stroke.* PhD (British Library R9260), Department of Cybernetics, University of Reading.

Anderson, A. H., Bard, E., Sotillo, C., Doherty-Sneddon, G., & Newlands, A. (1997). Limited visual control of the intelligibility of speech in face-to-face dialogue. *Perception and Psychophysics, 59*(4), pp. 580-592.

Anderson, A. H., Smallwood, L., MacDonald, R., Mullin, J., & Fleming, A. (2000). Video data and video links in mediated communication: what do users value? *International Journal of Human-Computer Studies, 52*(1), pp. 165-187.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of Learning Sciences, 4*, pp. 167-207.

Anthes, C. (2002). *Event and latency measurement in collaborative Virtual worlds*: The University of Reading.

Antunes, P., Guimarães, N., Segovia, J., & Cardenosa, J. (1995). *Beyond Formal Processes: Augmenting Workflow with Group Interaction Techniques.* In Proceedings of COCS '95: Proceedings of Conference on Organizational Computing Systems, New York, NY, USA, pp. 1-9.

Argyle, M. (1988). *Bodily Communication*: Routledge,an imprint of Taylor & Francis Books Ltd.

Argyle, M., & Cook, M. (1976). *Gaze and Mutual Gaze*: Cambridge University Press.

Argyle, M., & Graham, J. A. (1976). The Central Europe Experiment: Looking at Persons and Looking at Objects. *Environmental Psychology and Nonverbal Behavior, 1*(1), pp. 6–16.

Arthur, K. (1996). *Effects of field of view on task performance with head-mounted displays.* In Proceedings of Human Factors in Computing Systems, Vancouver, Canada, pp. 29-30.

Axelsson, A.-S., Abelin, A., Heldal, I., Nilsson, A., Schroeder, R., & Wideström, J. (1999). *Collaboration and Communication in Multi-User Virtual Environments: A Comparison of Desktop and Immersive Virtual Reality Systems for Molecular Visualization.* In Proceedings of Sixth UKVRSIG Conference, Salford University, pp. 107-117.

Axelsson, A.-S., 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), pp. 279-286.

Bailenson, J. N., Aharoni E. Beall, A. C., Guadagno, R. E., Dimov, A., & Blascovich, J. (2004). *Comparing Behavioral and Self-Report Measures of Embodied Agents Social Presence in Immersive Virtual Environments.* In Proceedings of the 7th Annual International Workshop on PRESENCE, Valencia, Spain.

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2001). Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments. *Presence: Teleoperators and Virtual Environments, 10*(6), pp. 583-598.

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003). Interpersonal Distance in Immersive Virtual Environments. *Personality and Social Psychology Bulletin, 29*(7), pp. 819-833.

Baker, H. H., Bhatti, N., Tanguay, D., Sobel, I., Gelb, D., Goss, M. E., MacCormick, J., Yuasa, K., Culbertson, B. W., & Malzbender, T. (2003). *Computation and*

Performance Issues in Coliseum: An Immersive Videoconferencing System. In Proceedings of MULTIMEDIA '03: Proceedings of the eleventh ACM international conference on Multimedia, New York, NY, USA, pp. 470-479.

Baños, R. M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., & Rey, B. (2004). Immersion and Emotion: Their Impact on the Sense of Presence. *CyberPsychology and Behavior, 7*(6), pp. 734-741.

Barakony, I., Frieb, W., & Schmalstieg, D. (2003). *Augmented Reality Videoconferencing for Collaborative Work.* In Proceedings of Hungarian Conference on Computer Graphics and Geomtery, Budapest, Hungary.

Barna, L. M. (1985). Stumbling Blocks in Intercultural Communication. In *Intercultural Communication: A Reader* (pp. 330-338): Wadsworth Publishing.

Barré, R., Pastoor, S., Conomis, C., Przewozny, D., Renault, S., Stachel, O., Duckstein, B., & Schenke, K. (2005). *Natural interaction in a desktop mixed reality environment.* In Proceedings of Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2005), Montreux (CH).

Basdogan, C., Ho, C.-H., Srinivasan, M. A., & Slater, M. (2000). An experimental study on the role of touch in shared virtual environments. *ACM Transactions on Computer-Human Interaction (ToCHI), 7*(4), pp. 443-460.

Becker, B., & Mark, G. (1998). *Social Conventions in Collaborative Virtual Environments.* In Proceedings of Collaborative Virtual Environments CVE'98, Manchester, pp. 47-55.

Bejczy, A. K., & Kim, W. S. (1990). *Predictive displays and shared compliance control for time-delayedtelemanipulation.* In Proceedings of Intelligent Robots and Systems '90. 'Towards a New Frontier of Applications', pp. 407-412.

Benford, S., Bowers, J., Fahlén, L. E., Greenhalgh, C., & Snowdon, D. (1995). *User Embodiment in Collaborative Virtual Environments.* In Proceedings of CHI'95: Proceedings of the ACM Conference on Human Factors in Computing Systems, pp. 242-249.

Benford, S., & Fahlén, L. E. (1993). *A Spatial Model of Interaction in Large Virtual Environments.* In Proceedings of 3rd European Conference on Computer Supported Cooperative Working, Millan, Italy.

Benford, S., Greenhalgh, C., Reynard, G., Brown, C., & Koleva, B. (1998). Understanding and Constructing Shared Spaces with Mixed-reality Boundaries. *ACM Transactions on Computer-Human Interaction (ToCHI), 5*(3), pp. 185-223.

Benford, S., Snowdon, D., Colebourne, A., O'Brien, J., & Rodden, T. (1997). *Informing the Design of Collaborative Virtual Environments.* In Proceedings of GROUP'97: International Conference on Supporting Group Work, pp. 71-80.

Bente, G., & Kraemer, N. C. (2002). *Virtual gestures: analyzing social presence effects of computer-mediated and computer-generated nonverbal behaviour.* In Proceedings of Fifth Annual International Workshop PRESENCE 2002, Porto, pp. 233-244.

Berlage, T., & Sohlenkamp, M. (1999). Visualizing Common Artefacts to Support Awareness in Computer-Mediated Cooperation. *Computer Supported Cooperative Work, 8*(3), pp. 207-238.

Bierbaum, A. (2001). *VR Juggler: A virtual Platform For Virtual Reality Application Development.* In Proceedings of IEEE VR 2001, Yokohama.

Billinghurst, M., & Kato, H. (2002). Collaborative augmented reality. *Communications of the ACM, 45*(7), pp. 64-70.

Biocca, F. (1997). The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments. *Journal of Computer-Mediated Communication (JCMC), 3*(2), pp. 12-26.

Bly, S. A. (1988). *A use of drawing surfaces in different collaborative settings.* In Proceedings of CSCW '88: Proceedings of the 1988 ACM conference on Computer-supported cooperative work, New York, NY, USA, pp. 250-256.

Bos, N., Olson, J., Gergle, D., Olson, G., & Wright, Z. (2002). *Effects of four computer-mediated communications channels on trust development.* In Proceedings of CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, USA, pp. 135-140.

Bos, N., Shami, S. N., Olson, J. S., Cheshin, A., & Nan, N. (2004). *In-group/out-group effects in distributed teams: an experimental simulation.* In Proceedings of CSCW '04: Proceedings of the 2004 ACM conference on Computer supported cooperative work, New York, NY, USA, pp. 429-436.

Bouguila, L., Ishii, M., & Sato, M. (2002). *Realizing a new step-in-place locomotion interface for virtual environment with large display system.* In Proceedings of EGVE '02: Proceedings of the workshop on Virtual environments 2002, Aire-la-Ville, Switzerland, Switzerland, pp. 197-207.

Bowers, J., O'Brien, J., & Pyckock, J. (1996). *Talk and embodiment in collaborative virtual environments.* In Proceedings of CHI'96, New York, pp. 58-65.

Bowman, D., Datey, A., Ryu, Y., Farooq, U., & Vasnaik, O. (2002a). *Empirical Comparison of Human Behavior and Performance with Different Display Devices for Virtual Environments.* In Proceedings of Human Factors and Ergonomics Society Annual Meeting, pp. 2134-2138.

Bowman, D. A., Gabbard, J. L., & Hix, D. (2002b). A Survey of Usability Evaluation in Virtual Environments: Classifcation and Comparison of Methods. *Presence: Teleoperators and Virtual Environments, 11*(4), pp. 404-424.

Bowman, D. A., Johnson, D. B., & Hodges, L. F. (2001). Testbed Evaluation of Virtual Environment Interaction Techniques. *Presence: Teleoperators and Virtual Environments, 10*(1), pp. 75-95.

Boyer, D. G., Handel, M. J., & Herbsleb, J. (1998). Virtual Community Presence Awareness. *SIGGROUP Bulletin, 19*(3), pp. 11-14.

Bradner, E., & Mark, G. (2002). *Why distance matters: effects on cooperation, persuasion and deception.* In Proceedings of CSCW '02: Proceedings of the 2002 ACM conference on Computer supported cooperative work, New York, NY, USA, pp. 226-235.

Brodlie, K., Duce, D. A., Gallop, J. R., Walton, J. P. R. B., & Wood, J. (2004). Distributed and Collaborative Visualization. *Computer Graphics Forum, 23*(2), pp. 223-251.

Brogni, A., Slater, M., & Steed, A. (2003a). *More breaks less presence.* In Proceedings.

Brogni, A., Slater, M., & Steed, A. (2003b). *More Breaks Less Presence.* In Proceedings of Presence 2003, The 6th Annual International Workshop on Presence, Aalborg, Danmark.

Broll, W. (1995). *Interacting in Distributed Collaborative Virtual Environments.* In Proceedings of VRAIS '95: Proceedings of the Virtual Reality Annual International Symposium, Washington, DC, USA, pp. 148-155.

Broll, W. (1997). *Populating the Internet: Supporting Multiple Users and Shared Applications with VRML.* In Proceedings of Proceedings of the VRML97 Symposium, Monterey, Ca., pp. 87-94.

Brooks, F. P. (1999). Whats Real About Virtual Reality? *IEEE Computer Graphics and Applications, 19*(6), pp. 16-27.

Burdea, G., Zhuang, J., Roskos, E., Silver, D., & Langrana, N. A. (1992). A Portable Dextrous Master with Force Feedback. *Presence: Teleoperators and Virtual Environments, 1*(1), pp. 18-28.

Burdea, G. C. (2000). Haptic Feedback for Virtual Reality. *special issue on Virtual Prototyping, International Journal of Design and Innovation Research, 2*(1), pp. 17-29.

Burgoon, M., Hunsaker, F. G., & Dawson, E. J. (1994). *Human Communication*. London: SAGE Publications.

Buttolo, P., Oboe, R., & Hannaford, B. (1997). Architectures for shared haptic virtual environments. *Computers and Graphics, 21*(4), pp. 421-429.

Buxton, W. A. S. (1992). *Telepresence: Integrating Shared Task and Person Spaces.* In Proceedings of Proceedings of the Conference on Graphics Interface '92, San Francisco, CA, USA, pp. 123-129.

Byrne, C. M. (1996). *Water on Tap: The Use of Virtual Reality as an Educational Tool.* University of Washington.

Bystrom, K.-E., & Barfield, W. (1999). Collaborative Task Performance for Learning Using a Virtual Environment. *Presence: Teleoperators and Virtual Environments, 8*(4), pp. 435-448.

Carlsson, C., & Hagsand, O. (1993). DIVE - A platform for multi-user virtual environments. *Computers and Graphics, 17*(6), pp. 663-669.

Casanueva, J. (2001). *Presence and Co-Presence in Collaborative Virtual Environments.* MSc, University of Cape Town, Cape Town, South Africa.

Casanueva, J., & Blake, E. (2000). *The Effects of Group Collaboration on Presence in a Collaborative Virtual Environment.* In Proceedings of Sixth Eurographics Workshop on Virtual Environments, EGVE'00, pp. 85-94.

Childers, L., Disz, T., Olson, R., Papka, M. E., Stevens, R., & Udeshi, T. (2000). *Access Grid: Immersive Group-to-Group Collaborative Visualization.* In Proceedings of Proceedings of the Fourth International Immersive Projection Technology Workshop.

Choi, H. R., Choi, B. H., & Ryew, S. M. (1997). *Haptic Display in the Virtual Collaborative Workspace Shared by Multiple Users.* In Proceedings of 6th IEEE Int. Workshop on Robot and Human Communication, pp. 478-483.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In *Perspectives on socially shared cognition*: WashingtonDC, APA Press.

Coomans, M. K. D., & Timmermans, H. J. P. (1998). *A VR-User Interface for Design by Features.* In Proceedings of the 4th Conference on Design and Decision Support Systems in Architecture and Urban Planning, DDSS'98, Maastricht, The Netherlands.

Cooper, R. (1974). The control of eye fixation by the meaning of spoken language. *Cognitive Psychology, 6*(1), pp. 84–107.

Crabtree, A. (2003). *Designing Collaborative Systems : A Practical Guide to Ethnography*: Springer.

Cruz-Neira, C., Sandin, D., DeFanti, T., Kenyon, R., & Hart, J. (1992). The CAVE: Audio Visual Experience Automatic Virtual Environment. *Communications of the ACM, 35*(6), pp. 65-72.

Darken, R. P., Bernatovich, D., Lawson, J. P., & Peterson, B. (1999). Quantitative Measures of Presence in Virtual Environments: The Roles of Attention and Spatial Comprehension. *CyberPsychology and Behavior, 2*(4), pp. 337-347.

Darn, S. (2005). Aspects of Nonverbal Communication. *The Internet TESL Journal, XI*(2).

Davis, T. R. V. (1984). The Influence of the Physical Environment in Offices. *Academy of Management Review, 9*(2), pp. 271-283.

Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1995). The evolution of research on collaborative learning. In *Learning in Humans and Machine: Towards an interdisciplinary learning science* (pp. 189-211): Oxford: Elsevier.

Draper, J. V., Kaber, D. B., & Usher, J. M. (1998). Telepresence. *Human Factors, 40*(3), pp. 354-375.

Durlach, N., & Slater, M. (2000). Presence in Shared Virtual Environments and Virtual Togetherness. *Presence: Teleoperators and Virtual Environments, 9*(2), pp. 214-217.

Ekman, P., & Friesan, W. (1969). The repertoire of nonverbal behavior: Categories,. *Semiotica, 1*, pp. 49-98.

Ellis, C. A., Gibbs, S. J., & Rein, G. (1991). Groupware: Some Issues and Experiences. *Communications of the ACM, 34*(1), pp. 39-58.

Ellis, S. R. (1996). Virtual Environments and Environmental Instruments. In K. Carr & R. England (Eds.), *Simulated and Virtual Realities - Elements of Perception* (pp. 11-51): London: Taylor and Francis.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1), pp. 32-64.

Flach, J. M. (1998). The Reality of Experience: Gibson's Way. *Presence: Teleoperators and Virtual Environments, 7*(1), pp. 90-95.

Floyd, S., Jacobson, V., McCanne, S., Liu, C.-G., & Zhang, L. (1995). *A Reliable Multicast Framework for Light-weight Sessions and Application Level Framing.* In Proceedings of SIGCOMM'95: Applications, Technologies, Architectures,and Protocols for Computer Communication, Cambridge, Massachusetts, pp. 342-356.

Forgas, J. P. (1985). *Interpersonal Behaviour: Psychology of Social Interaction*: Elsevier.

Fraser, M., Benford, S., Hindmarsh, J., & Heath, C. (1999). *Supporting Awareness and Interaction through Collaborative Virtual Interfaces.* In Proceedings of UIST'99: 12th Annual ACM Symposium on User Interface Software and Technology, New York, NY, USA, pp. 27-36.

Frécon, E. (2004). DIVE: communication architecture and programming model. *IEEE Communications Magazine, Special Issue on Networked Virtual Environments, 42*(4), pp. 34-40.

Frécon, E., Greenhalgh, C., & Stenius, M. (1999). *The DiveBone - An Application-Level Network Architecture for Internet-based CVEs.* In Proceedings of VRST'99: Symposium on Virtual Reality Software and Technology 1999, University College London, UK, pp. 58-65.

Frécon, E., Smith, G., Steed, A., Stenius, M. a., & Ståahl, O. (2001). An Overview of the COVEN Platform. *Presence: Teleoperators and Virtual Environments, 10*(1), pp. 109-127.

Fussell, S. R., Kraut, R. E., & Siegel, J. (2000). *Coordination of Communication: Effects of Shared Visual Context on Collaborative Work.* In Proceedings of CSCW '00: Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, New York, NY, USA, pp. 21-30.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin Co.

Goleman, D. (1996). *Emotional Intelligence: Why It Can Matter More Than IQ*: Bloomsbury.

Gonzalez, D. A. (2005). *AGJuggler: An architecture for virtual reality within a collaboration environment.* Purdue University, West Lafayette, IN 47907, USA.

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*: New York: Academic.

Greenhalgh, C. M. (1999). *Analysing Awareness Management in Distributed Virtual Environments.* In Proceedings of 2nd Annual Workshop on System Aspects of Sharing a Virtual Reality, part of CVE '98, Manchester, UK.

Greenhalgh, C. M., & Benford, S. D. (1995). MASSIVE: A Virtual Reality System for Tele-conferencing. *ACM Transactions on Computer-Human Interaction (ToCHI), 2*(3), pp. 239-261.

Greenhalgh, C. M., Bullock, A., Fréon, E., Lloyd, D., & Steed, A. (2001). Making Networked Virtual Environments Work. *Presence: Teleoperators and Virtual Environments, 10*(2), pp. 142-159.

Hall, E. T. (1963). A System for the notation of proxemic behavior. *American Anthropologist, 65*(5), pp. 1003.

Hall, E. T. (1968). Proxemics. *Current Anthropology, 9*(2), pp. 83.

Hardman, V., Sasse, A., Handley, M., & Watson, A. (1995). *Reliable Audio for Use over the Internet.* In Proceedings of INET'95, Honolulu, Hawaii.

Heldal, I., Steed, A., Spante, M., Schroeder, R., Bengtsson, S., & Partanan, M. (2005). Successes and Failures in Co-Present Situations. *Presence: Teleoperators and Virtual Environments, 14*(5).

Hindmarsh, J., Fraser, M., Heath, C., & Benford, S. (2001). Collaborative Virtual Environments. In *Virtually Missing the Point: Configuring CVEs for Object-Focused Interaction* (pp. 115-139). London, UK: Springer Verlag.

Hindmarsh, J., Fraser, M., Heath, C., Benford, S., & Greenhalgh, C. (2000). Object-Focused Interaction in Collaborative Virtual Environments. *ACM Transactions on Computer-Human Interaction (ToCHI), 7*(4), pp. 477-509.

Hix, D., & Hartson, R. H. (1993). *Developing User Interfaces : Ensuring Usability Through Product and Process*: John Wiley and Sons.

Hollan, J., & Stornetta, S. (1992). *Beyond being There.* In Proceedings of CHI '92: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, pp. 119-125.

Holmes, J. G., & Rempel, J. K. (1989). Trust in close relationships. *Review of personality and social psychology: Close relationships, 10*, pp. 187-219.

Huettig, F., & Altmann, G. T. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition, 96*(1), pp. B23-B32.

Hurmuzlu, Y., Ephanov, A., & Stoianovici, D. (1998). Effect of a Pneumatically Driven Haptic Interface on the Perceptional Capabilities of Human Operators. *Presence - Teleoperators and Virtual Environments, 7*(3), pp. 290–307.

Hutchby, I. (2001). *Conversation and Technology*: Cambridge University Press.

Hutchby, I., & Wooffitt, R. (1998). *Conversation analysis*: Cambridge University Press.

IJsselsteijn, W. A., Freeman, J., & Ridder, H. (2001). Presence: Where Are We? *Cyberpsychology and Behavior, 4*(2), pp. 179–182.

Johnson, A., Leigh, J., DeFanti, T., Brown, M., & Sandin, D. (1998). *CAVERN: the CAVE Research Network.* In Proceedings of Proceedings of 1st International Symposium on Multimedia Virtual Laboratory, Tokyo, Japan, pp. 15-27.

Joiner, J. T. E. (1994). Contagious depression: existence, specificity to depressed symptoms, and the role of reassurance seeking. *Journal of Personality and Social Psychology, 67*(2), pp. 287-296.

Jorgensen, C., Ogden, J., Willis, K., Blessing, M., Caudell, K. A., Patrick, G., & Caudell, T. P. (1997). *Locomotion in a virtual environment: Performance measures and* *physiological responses.* In Proceedings of Human Factors and Ergonomics Society 41st Annual Meeting, Albuquerque, New Mexico, pp. 1148-1151.

Jung, J. Y., Adelstein, B. D., & Ellis, S. R. (2000). *Discriminability of prediction artifacts in a time delayed virtual environment.* In Proceedings of the 44th Annual Human Factors and Ergonomics Society Meeting, pp. 499-502.

Kalawsky, R. S. (2000). *The Validity of Presence as a Reliable Human Performance Metric in Immersive Environments.* In Proceedings of 3rd International Workshop on Presence, Delft, Netherlands, pp. 1-16.

Kauff, P., & O.Schreer. (2002). *Virtual Team User Environments - A Step From Tele-Cubicles Towards Distributed Tele-Colaboration in Mediated Workspaces.* In Proceedings of Int. IEEE Conf. on Multimedia and Expo (ICME 2002), Lausanne, Switzerland.

Kendon, A. (1990). *Conducting Interaction : Patterns of Behavior in Focused Encounters (Studies in Interactional Sociolinguistics)*: Cambridge University Press.

Kim, J., Kim, H., Tay, B. K., Muniyandi, M., Srinivasan, M. A., Jordan, J., Mortensen, J., Oliveira, M., & Slater, M. (2004). Transatlantic touch: a study of haptic collaboration over long distance. *Presence: Teleoperators and Virtual Environments, 13*(3), pp. 328.

Kjeldskov, J. (2001). *Interaction: Full and partial Immersive Virtual Reality Displays.* In Proceedings of IRIS24, University of Bergen, Bergen, pp. 587-600.

Knapp, M. L. (1978). *Nonverbal Communication in Human Interaction*: Holt, Rinehart and Winston.

Kraut, R. E., Fussell, S. R., & Siegel, J. (2003). Visual Information as a Conversational Resource in Collaborative Physical Tasks. *Human-Computer Interaction, 18*(1/2), pp. 13-49.

Lanier, J. (2001). Virtually There. *Scientific American, 284*(4), pp. 66-75.

Lapointe, J.-F., & Vinson, N. G. (2002). *Effects of Joystick Mapping and Field-of-View on Human Performance in Virtual Walkthroughs.* In Proceedings of the 1st International Symposium on 3D Data Processing Visualization and Transmission, Padova, Italy, pp. 490-493.

Leigh, J., Johnson, A., & DeFanti, T. (1999). *A Review of Tele-Immersive Applications in the CAVE Research Network.* In Proceedings of IEEE VR `99, Houston, TX, pp. 180-187.

Lin, J. J.-W., Duh, H. B. L., Abi-Rached, H., Parker, D. E., & Iii, T. F. A. (2002). *Effects of Field of View on Presence, Enjoyment, Memory, and Simulator Sickness in a Virtual Environment.* In Proceedings of VR '02: Proceedings of the IEEE Virtual Reality Conference 2002, Washington, DC, USA, pp. 164.

Linebarger, J. M., Janneck, C. D., & Kessler, D. G. (2003). *Shared Simple Virtual Environment: An Object-Oriented Framework for Highly Interactive Group Collaboration.* In Proceedings of DS-RT '03: Proceedings of the Seventh IEEE International Symposium on Distributed Simulation and Real-Time Applications, Washington, DC, USA, pp. 170-180.

Linebarger, J. M., & Kessler, D. (2004). Concurrency Control Mechanisms for Closely Coupled Collaboration in Multithreaded Peer-to-Peer Virtual Environments. *Presence: Teleoperators and Virtual Environments, 13*(3), pp. 296-314.

Lok, B., Naik, S., Whitton, M., & Frederick P. Brooks, J. (2003). Effects of Handling Real Objects and Self-Avatar Fidelity on Cognitive Task Performance and Sense of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments, 12*(6), pp. 615-628.

Loureiro, R., Amirabdollahian, F., Topping, M., Driessen, B., & Harwin, W. (2005). Upper Limb Mediated Stroke Therapy – GENTLE/s Approach. *In Special Issue on Rehabilitation Robotics, Journal of Autonomous Robots, 15*(1), pp. 35-51.

Macedonia, M. R., Zyda, M. J., Pratt, D. R., Barham, P. T., & Zeswitz, S. (1994). NPSNET: A Network Software Architecture for Large-Scale Virtual Environment. *Presence: Teleoperators and Virtual Environments, 3*(4), pp. 265-287.

Manek, D. (2004). *Effects of Visual Displays on 3D Interaction in Virtual Environments.* Virginia Polytechnic Institute and State University.

Marcelino, L., Murray, N., & Fernando, T. (2003). A Constraint Manager to Support Virtual Maintainability. *Computers and Graphics, 27*, pp. 19-26.

Margery, D., Arnaldi, B., & Plouzeau, N. (1999). A General Framework for Cooperative Manipulation in Virtual Environments. *Virtual Environments '99*, pp. 169-178.

Mark, G., Grudin, J., & Poltrock, S. E. (1999). *Meeting at the Desktop: An Empirical Study of Virtually Collocated Teams.* In Proceedings of ECSCW'99: Sixth European Conference on Computer Supported Cooperateive Work.

Markus, L. M., & Connolly, T. (1990). *Why CSCW Applications fail: Problems in the Adoption of Interdependent Work Tools.* In Proceedings of CSCW '90: Proceedings of the 1990 ACM Conference on Computer Supported Cooperative Work, New York, NY, USA, pp. 371-380.

Massie, T. H. (1993). *Initial Haptic Explorations with the Phantom: Virtual Touch Through Point Interaction.* Massachusetts Institute of Technology.

Meehan, M., Insko, B., Whitton, M., & Brooks, F. P. (2001). *Physiological Measures of Presence in Virtual Environments.* In Proceedings of 4rd International Workshop on Presence, Philadelphia, USA.

Meehan, M., Razzaque, S., Whitton, M. C., & Brooks, J. F. P. (2003). *Effect of Latency on Presence in Stressful Virtual Environments.* In Proceedings of IEEE Virtual Reality 2003, Los Angeles, CA, pp. 141-148.

Milgram, P., & Fumio, K. (1994). A Taxonomy of Mixed Reality Displays. *IEICE Transactions on Information Systems, E77-D*(12), pp. 1321-1329.

Mine, M. R., Frederick P. Brooks, J., & Sequin, C. H. (1997). *Moving objects in space: exploiting proprioception in virtual-environment interaction.* In Proceedings of SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques, New York, NY, USA, pp. 19-26.

Molet, T., Aubel, A., Capin, T., Carion, S. e., Lee, E., Magnenat-Thalmann, N., Noser, H., Pandzic, I., Sannier, G. e., & Thalmann, D. (1999). Anyone for Tennis? *Presence: Teleoperators and Virtual Environments, 8*(2), pp. 140-156.

Mortensen, J., Vinayagamoorthy, V., Slater, M., Steed, A., Lok, B., & Whitton, M. C. (2002). *Collaboration in tele-immersive environments.* In Proceedings of Eighth Eurographics workshop on Virtual environments, Barcelona, Spain, pp. 93-101.

Nichols, S., Haldane, C., & Wilson, J. R. (2000). Measurement of presence and its consequences in virtual environments. *International Journal of Human Computer Studies, 52*(3), pp. 471.

Nilsson, A., Heldel, I., Axelsson, A.-S., & Schroeder, R. (2002). The long-term uses of shared virtual environments: an exploratory study. In *The social life of avatars: presence and interaction in shared virtual environments* (pp. 112-126). New York, NY, USA: Springer-Verlag New York, Inc.

O'Conaill, B., Whittaker, S., & Wilbur, S. (1993). Conversations Over Video Conferences: An Evaluation of the Spoken Aspects of Video-Mediated Communication. *Human-Computer Interaction, 8*(4), pp. 389-428.

Ohba, K., Kawabata, S., Chong, N. Y., Komoriya, K., Matsumaru, T., Matsuhira, N., Takase, K., & Tanie, K. (1999). *Remote Collaboration through Time Delay in Multiple Teleoperation.* In Proceedings of IEEE International Conference on Intelligent Robots and Systems(IROS99), pp. 1866-1871.

Olson, J. S., Olson, G. M., & Meader, D. K. (1995). *What mix of video and audio is useful for remote real-time work?* In Proceedings of CHI ’95, pp. 362-368.

Osawa, N. (2006). *Automatic adjustments for efficient and precise positioning and release of virtual objects.* In Proceedings of ACM international Conference on Virtual Reality Continuum and its Applications, VRCIA, HongKong, pp. 121-128.

Painter, S. L. (1991). Personal Space and Privacy: Implications for Correctional Institutions. *Institutional Design and Correctional Environments, 3*(2), pp. 1-8.

Park, K. S., Cho, Y. J., Krishnaprasad, N. K., Scharver, C., Lewis, M. J., Leigh, J., & Johnson, A. E. (2000). *CAVERNsoft G2: a toolkit for high performance tele-immersive collaboration.* In Proceedings of VRST '00: Proceedings of the ACM symposium on Virtual reality software and technology, New York, NY, USA, pp. 8-15.

Pausch, R., Crea, T., & Conway, M. (1992). A literature survey for virtual environments: military flight simulator visual systems and simulator sickness. *Presence: Teleoper. Virtual Environ., 1*(3), pp. 344-363.

Pausch, R., Proffitt, D., & Williams, G. (1997). *Quantifying Immersion in Virtual Reality.* In Proceedings of SIGGRAPH '97: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, New York, NY, USA, pp. 13-18.

Pearson, J. C., & Nelson, P. E. (1994). *Understanding and Sharing: An Introduction to Speech Communication*: Brown and Benchmark Pub.

Pedersen, E. R., McCall, K., Moran, T. P., & Halasz, F. G. (1993). *Tivoli: an electronic whiteboard for informal workgroup meetings.* In Proceedings of CHI '93: Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, USA, pp. 391-398.

Penington, D. L. (1985). Intercultural Communication. In *Intercultural Communication: A Reader* (pp. 31-36): Wadsworth Publishing.

Poupyrev, I., Weghorst, S., Billinghurst, M., & Ichikawa, T. (1998). Egocentric object manipulation in virtual environments: empirical evaluation of interaction techniques. *Computer Graphics Forum, 17*(3), pp. 41-52.

Prasolova-Førland, E. (2005). *Analyzing Place Metaphors in Different Educational CVE's: Lessons Learned.* In Proceedings of IADIS International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2005), Porto, Portugal, pp. 325-332.

Psotka, J. (1995). Immersive training systems: Virtual reality and education and training. *Instructional Science, 23*(5-6), pp. 405-431.

Purbrick, J., & Greenhalgh, C. (2003). An extensible event-based infrastructure for networked virtual worlds. *Presence: Teleoperators and Virtual Environments, 12*(1), pp. 68-84.

Raskar, R., Welch, G., Cutts, M., Lake, A., Stesin, L., & Fuchs, H. (1998). *The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays.* In Proceedings of SIGGRAPH '98: Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, New York, NY, USA, pp. 179-188.

Riva, G. (1999). Virtual Reality as Communication Tool: A Sociocognitive Analysis. *Presence: Teleoperators and Virtual Environments, 8*(4), pp. 462-468.

Roberts, D., Al-Liabi, M., Wolff, R., Otto, O., & Al-Khalifah, A. (2005a). *Reducing Fragmentation in Telecollaboration by Using IPT Interfaces.* In Proceedings of IPT-EGVE Workshop, pp. 211-216.

Roberts, D., Heldal, I., Otto, O., & Wolff, R. (2006). Factors influencing creativity in collaborative virtual environments. *To be published in VR Journal special issue: Collaborative Virtual Environments for Creative People*.

Roberts, D., & Wolff, R. (2004a). *Controlling Consistency within Collaborative Virtual Environments.* In Proceedings of The Eighth IEEE International Symposium on Distributed Simulation and Real Time Applications, DS-RT 2004, pp. 46-52.

Roberts, D., Wolff, R., & Otto, O. (2003). Constructing a Gazebo: Supporting team work in a tightly coupled, distributed task in virtual reality. *Presence: Teleoperators and Virtual Environments, 12*(6), pp. 644-668.

Roberts, D., Wolff, R., & Otto, O. (2005b). Supporting a Closely Coupled Task between a Distributed Team: Using Immersive Virtual Reality Technology. *Computing and Informatics, 24*, pp. 1001-1023.

Roberts, D., Wolff, R., Otto, O., Kranzlmueller, D., Anthes, C., & Steed, A. (2004b). *Supporting Social Human Communication between Distributed Walk-in Displays.* In Proceedings of VRST 2004, pp. 81-88.

Roberts, D. J. (2003). Communication Infrastructures for Inhabited Information Spaces. *Inhabited Information Systems*, pp. 233-268.

Roberts, D. J., Sharkey, P., & Sandoz, P. (1996). *A Real-time Predictive Architecture for Distributed Virtual Reality.* In Proceedings of 1st ACM SIGGRAPH Workshop on Simulation & Interaction in Virtual Environments, pp. 279-288.

Roberts, D. J., Strassner, J., Worthington, B. G., & Sharkey, P. (1999). *Influence of the Supporting Protocol on the Latencies Induced by Concurrency Control within a Large Scale Multi User Distributed Virtual Reality System.* In Proceedings of Interanational Conference on Virtual Worlds and Simulation (VWSIM), SCS Western Multi-conference '99, San Francisco, CA,, Vol. *31*, pp. 70-75.

Robinson, M., Pekkola, S., Korhonen, J., Hujala, S., Toivonen, T., & Saarinen, M.-J. (2001). Extending the Limits of Collaborative Virtual Environments. In *Collaborative Virtual Environments. Digital Places and Spaces for Interaction* (pp. 21-42): Springer Verlag.

Rodden, T. (1996). *Populating the Application: A Model of Awareness for Cooperative Applications.* In Proceedings of CSCW '96: Proceedings of the 1996 ACM conference on Computer Supported Cooperative Work, New York, NY, USA, pp. 87-96.

Ruddle, R. A., Savage, J. C., & Jones, D. M. (2002). Symmetric and Asymmetric Action Integration during Cooperative object Manipulation in Virtual Environments. *ACM Transactions on Computer-Human Interaction (ToCHI), 9*(4), pp. 285-308.

Salem, B. (2005). Commedia Virtuale: from theatre to avatars. *Digital Creativity Journal, 16*(3), pp. 129-139.

Sallnäs, E.-L. (2005). Effects of Communication Mode on Social Presence, Virtual Presence, and Performance in Collaborative Virtual Environments. *Presence - Teleoperators and Virtual Environments, 14*(4), pp. 434-449.

Sallnäs, E.-L., Rassmus-Gröhn, K., & Sjöström, C. (2000). Supporting Presence in Collaborative Environments by Haptic Force Feedback. *ACM Transactions on Computer-Human Interaction (ToCHI), 7*(4), pp. 461.

Samover, L. A., & Porter, R. E. (1985). Approaching Intercultural Communication. In *Intercultural Communication: A Reader* (pp. 15-30): Wadsworth Publishing.

Sander, I. (2005). *Studying of Rehabilitation for Patients with Balance Disorders and Loss of Motor Coordination in Real and Virtual Environments.* MSc, CVE, University of Salford, Mancherster.

Sas, C., & O'Hare, G. M. P. (2003). Presence equation: an investigation into cognitive factors underlying presence. *Presence: Teleoperators and Virtual Environments, 12*(5), pp. 523.

Schmalstieg, D., Fuhrmann, A., Hesina, G., Szalav\'ari, Z., Encarnaccão, M. L., Gervautz, M., & Purgathofer, W. (2002). The Studierstube Augmented Reality Projects. *Presence: Teleoperators and Virtual Environments, 11*(1), pp. 33-54.

Schmidt, K. (2002). The Problem with 'Awareness': Introductory Remarks on 'Awareness in CSCW'. *Computer Supported Cooperative Work, 11*(3), pp. 285-298.

Schroeder, R. (2002). *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London, UK: Springer Verlag.

Schroeder, R., Steed, A., Axelsson, A. S., Heldal, I., Abelin, A., Widestroem, J., Nilsson, A., & Slater, M. (2001). Collaborating in networked immersive spaces: as good as being there together? *Computers and Graphics, 25*(5), pp. 781-788.

Schubert, T. W., Friedmann, F., & Regenbrecht, H. T. (1999). Embodied presence in virtual environments. In *In Visual Representations and Interpretations* (pp. 268-278): Springer-Verlag.

Schuemie, M. J., Straaten, P. v., M.Krijn, & Mast, C. A. P. G. v. (2001). Research on Presence in VR: a Survey. *CyberPsychology and Behavior, 4*(2), pp. 183-202.

Seelig, M., Roberts, D. J., Harwin, W., Otto, O., & Wolff, R. (2004). *A haptic interface for linked immersive and desktop displays: Maintaining sufficient frame rate for haptic rendering.* In Proceedings of 17th International Conference on Parallel and Distributed Computing Systems (PDCS 2004), San Francisco, CA, USA, pp. 478-483.

Sharkey, P. M., Ryan, M. D., & Roberts, D. J. (1998). *Theoretical and Practical Constraints on Shared Virtual Environments due to Latency.* In Proceedings of Systems Aspects of Sharing a Virtual Reality Workshop, Collaborative Virtual Environments (CVE 98), Manchester, UK.

Sheridan, T. B. (2000). *Interaction, Imagination and Immersion some Research Needs.* In Proceedings of VRST '00: Proceedings of the ACM symposium on Virtual reality software and technology, New York, NY, USA, pp. 1-7.

Short, J., Williams, E., & Christie, B. (1976). *The Social Psychology of Telecommunications*. New York: Wiley.

Singer, M. J., Ehrlich, J., & Cinq-Mars, S. (1995). *Task Performance in Virtual Environments: Stereoscopic Versus Monoscopic Displays and Head-Coupling*.

Singer, M. J., & Witmer, B. G. (1999). On Selecting the Right Yardstick. *Presence: Teleoperators and Virtual Environments, 8*(5), pp. 566-573.

Sitzman, K. (2003). Likert type questions may be helpful for baseline inquiry. *AAOHN Journal, 51*(7), pp. 320.

Skinner, B. F. (1971). Beyond Freedom and Dignity: Alfred A. Knopf.

Slater, M. (1999). Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire. *Presence: Teleoperators and Virtual Environments, 8*(5), pp. 560.

Slater, M., Sadagic, A., Usoh, M., & Schroeder, R. (2000a). Small Group Behaviour in a Virtual and Real Environment: A Comparative Study. *Presence: Teleoperators and Virtual Environments, 9*(1), pp. 37-51.

Slater, M., & Steed, A. (2000b). A Virtual Presence Counter. *Presence: Teleoperators and Virtual Environments, 9*(5), pp. 413-434.

Slater, M., & Usoh, M. (1993). *The influence of a virtual body on presence in immersive virtual environments.* In Proceedings of Virtual Reality International 93: Proceedings of the third annual conference on Virtual Reality, London, pp. 34-42.

Slater, M., Usoh, M., & Steed, A. (1994). Depth of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments, 3*(2), pp. 130-144.

Snow, M. P. (1996). *Charting presence in virtual environments and its effects on performance.* PhD, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Snow, M. P., & Williges, R. C. (1998). Empirical models based on free-modulus magnitude estimation of perceived presence in virtual environments. *Human Factors, 40*(3), pp. 386-402.

Snowdon, D., Churchill, E. F., & Munro, A. J. (2001). Collaborative Virtual Environments: Digital Spaces and Places for CSCW: An Introduction. In *Collaborative Virtual Environments: Digital Places and Spaces for Interaction* (pp. 3-17). London, UK: Springer Verlag.

Speeter, T. H. (1992). Transforming Human Hand Motion for Telemanipulation. *Presence: Teleoperators and Virtual Environments, 1*(1), pp. 63-79.

Ståhl, O. (1999). *Meetings for Real-Experiences from a Series of VR-based Project Meetings.* In Proceedings of VRST '99: Proceedings of the ACM Symposium on Virtual Reality Software and Technology, New York, NY, USA, pp. 164-165.

Stanney, K., & Salvendy, G. (1998). Aftereffects and Sense of Presence in Virtual Environments: Formulation of a Research and Development Agenda. *International Journal of Human-Computer Interaction, 10*(2), pp. 135-187.

Steed, A., Mortensen, J., & Frécon, E. (2001). Spelunking: Experiences using the DIVE System on CAVE-like Platforms. *Immersive Projection Technologies and Virtual Environments*, pp. 153-164.

Steed, A., & Parker, C. (2005). Evaluating Effectiveness of Interaction Techniques across Immersive Virtual Environment Systems. *Presence: Teleoperators and Virtual Environments, 14*(5), pp. (in press).

Steed, A., Slater, M., Sadagic, A., Tromp, J., & Bullock, A. (1999). *Leadership and collaboration in virtual environments.* In Proceedings of IEEE Virtual Reality, pp. 112-115.

Stevens, B., Jerrams-Smith, J., Heathcote, D., & Callear, D. (2002). Putting the virtual into reality: assessing object-presence with projection-augmented models. *Presence: Teleoperators and Virtual Environments, 11*(1), pp. 79-92.

Stone, R. J. (2000). *Haptic Feedback: A Brief History from Telepresence to Virtual Reality.* In Proceedings of Haptic Human-Computer Interaction, Glasgow, Scotland, pp. 1-7.

Tan, D. S., Gergle, D., Scupelli, P. G., & Pausch, R. (2003). *With Similar Visual Angles, Larger Displays Improve Spatial Performance.* In Proceedings of Conference on Human factors in Computing Systems (CHI), Ft. Lauderdale, Florida, USA*,* Vol. *5*, pp. 217-224.

Tannen, D. (2001). *You Just Don't Understand: Women and Men in Communication*: Harper Paperbacks.

Thie, S., & Wijk, J. (1998). *A General Theory on Presence.* In Proceedings of Presence in Shared Virtual Environments Workshop, First International Workshop on Presence.

Tickle-Degnen, L., & Puccinelli, N. M. (1999). The Nonverbal Expression of Negative Emotions: Peer and Supervisor Responses to Occupational Therapy Students'

Emotional Attributes. *Occupational Therapy Journal of Research, 19*(1), pp. 18-39.

Towles, H., Chen, W.-C., Yang, R., Kum, S.-U., Fuchs, H., Kelshikar, N., Mulligan, J., Daniilidis, K., Holden, L., Zeleznik, B., Sadagic, A., & Lanier, J. (2002). *3D Tele-Collaboration Over Internet2.* In Proceedings of International Workshop on Immersive Telepresence (ITP2002), Juan Les Pins, France.

Tromp, J., Bullock, A., Steed, A., Sadagic, A., Slater, M., & Frécon, E. (1998). Small Group Behavior Experiments in the Coven Project. *IEEE Comput. Graph. Appl., 18*(6), pp. 53.

Ulhaas, K. D., & Schmalstieg, D. (2001). *Finger Tracking for Interaction in Augmented Environments.* In Proceedings of ISAR '01: Proceedings of the IEEE and ACM International Symposium on Augmented Reality (ISAR'01), Washington, DC, USA, pp. 55.

Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Using Presence Questionnaires in Reality. *Presence: Teleoperators and Virtual Environments, 9*(5), pp. 497-503.

Van der Linde, R. Q., Lammertse, P., Frederikson, E., & Ruiter, B. (2002). *The Haptic Master, a new high-performance haptic interface.* In Proceedings of EUROHAPTICS 2002.

The VirtuSphere. (2004). from http://www.virtusphere.com

Wainfan, L., & Davis, P. (2004). *Challenges in Virtual Collaboration*: RAND National Security Research Division.

Welch, R., Blackmon, T. T., Liu, A., Mellers, B. A., & Stark, L. W. (1996). The Effects of Pictorial Realism, Delay of Visual Feedback, and Observer Interactivity on the Subjective Sense of Presence. *Presence: Teleoperators and Virtual Environments, 5*(3), pp. 263-273.

Welch, R. B. (1997). The Presence of Aftereffects. In *Designs of computing Systems: Cognitive Considerations* (pp. 271-276): Amsterdam: Elsevier.

Welch, R. B. (1999). How Can We Determine if the Sense of Presence Affects Task Performance? *Presence: Teleoperators and Virtual Environments, 8*(5), pp. 574-577.

Whittaker, S. (1995). Rethinking video as a technology for interpersonal communications: theory and design implications. *International Journal of Human-Computer Studies, 42*(5), pp. 501-529.

Whittaker, S. (2003). Things to talk about when talking about things. *Human-Computer Interaction, 18*(1-2), pp. 149-170.

Whittaker, S., Geelhoed, E., & Robinson, E. (1993). Shared workspaces: how do they work and when are they useful? *International Journal of Man-Machine Studies, 39*(5), pp. 813-842.

Wiederhold, B. K., Davis, R., & Wiederhold, M. D. (1998). The effects of immersiveness on physiology. In *Virtual Environments in Clinical Psychology and Neuroscience: Methods and Techniques in Advanced Patient-therapist Interaction* (pp. 52-60): IOS Press.

Wolff, R. (2005). *Reconfigurable Event Handling for CVE Systems*: Centre for Virtual Environments, University of Salford, UK.

Wolff, R. (2006). *Supporting Closely Coupled Collaboration around Shared Objects in Immersive Virtual Environments.* PhD, University of Salford, Salford.

Wolff, R., Roberts, D. J., & Otto, O. (2004). A Study of Event Traffic during the Shared Manipulation of Objects within Collaborative Virtual Environments. *Presence: Teleoperators and Virtual Environments, 13*(3), pp. 251-262.

Wood, J. T. (2004). *Gendered Lives: Communication, Gender, and Culture*: Wadsworth Publishing.

Yang, R., Nashel, A., & Towles, H. (2004). *Interactive 3D Teleconferencing with User-Adaptive Views.* In Proceedings of ETP '04: Proceedings of the 2004 ACM SIGMM Workshop on Effective Telepresence, New York, NY, USA, pp. 50-51.

Youngblut, C. (2003). *Experience of Presence in Virtual Environments* (Tech. Report No. D-2960): Institute for Defense Analyses (ida.org).

Zimmons, P., & Panter, A. (2003). *The Influence of Rendering Quality on Presence And Task Performance in a Virtual Environment.* In Proceedings of VR '03: Proceedings of the IEEE Virtual Reality 2003, Washington, DC, USA, pp. 293.