



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/4965/>

Article:

Jones, D.M., Ruddle, R.A. and Savage, J.C. (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. ISSN: 1073-0516

<https://doi.org/10.1145/586081.586084>

Reuse

See Attached

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

promoting access to White Rose research papers



Universities of Leeds, Sheffield and York
<http://eprints.whiterose.ac.uk/>

This is an author produced version of a paper published in **ACM Transactions on Computer-Human Interaction**.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/4965/>

Published paper

Jones, D.M., Ruddle, R.A. and Savage, J.C. (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.
<http://dx.doi.org/10.1145/586081.586084>

Symmetric and Asymmetric Action Integration During Cooperative Object Manipulation in Virtual Environments

ROY A. RUDDLE

University of Leeds

JUSTIN C. D. SAVAGE AND DYLAN M. JONES

Cardiff University

Cooperation between multiple users in a virtual environment (VE) can take place at one of three levels. These are defined as where users can perceive each other (Level 1), individually change the scene (Level 2), or simultaneously act on and manipulate the same object (Level 3). Despite representing the highest level of cooperation, multi-user object manipulation has rarely been studied. This paper describes a behavioral experiment in which the *piano movers' problem* (maneuvering a large object through a restricted space) was used to investigate object manipulation by pairs of participants in a VE. Participants' interactions with the object were integrated together either symmetrically or asymmetrically. The former only allowed the common component of participants' actions to take place, but the latter used the mean. Symmetric action integration was superior for sections of the task when both participants had to perform similar actions, but if participants had to move in different ways (e.g., one maneuvering themselves through a narrow opening while the other traveled down a wide corridor) then asymmetric integration was superior. With both forms of integration, the extent to which participants coordinated their actions was poor and this led to a substantial *cooperation overhead* (the reduction in performance caused by having to cooperate with another person).

Categories and Subject Descriptors: I.3.6 [Computer Graphics]: Methodology and Techniques - Interaction Techniques. I.3.6 [Computer Graphics]: Three-Dimensional Graphics and Realism - Virtual Reality. H.5.2 [Information Interfaces and Presentation]: User Interfaces - Input devices and strategies.

General Terms: Experimentation, Human Factors, Measurement, Performance.

Additional Key Words and Phrases: Virtual environments, object manipulation, piano movers' problem, rules of interaction.

This research was supported by the UK Engineering and Physical Sciences Research Council.

Authors' addresses: R. A. Ruddle, School of Computing, University of Leeds, Leeds, UK, LS2 9JT; email: roy@comp.leeds.ac.uk; J. C. D. Savage and D. M. Jones, School of Psychology, Cardiff University, Cardiff, UK, CF10 3YG; email: savagejc@cf.ac.uk; jonesdm@cf.ac.uk.

Permission to make digital/hard copy of part of this work for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication, and its date of appear, and notice is given that copying is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee.

© 2001 ACM 1073-0516/01/0300-0034 \$5.00

1. INTRODUCTION

This paper brings together two well-known themes in virtual environment (VE) research, those of object manipulation, and interaction in collaborative virtual environments (CVEs). Object manipulation is one of the primary types of task that is performed in a VE. Research in this area has focused on single-user interaction, making detailed studies of the manner in which the separate components of manipulation are performed (translation and rotation), the degree to which those components are coordinated, comparisons between different interface metaphors, and the similarity of virtual to real-

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

world object manipulation [e.g., Bowman et al., 2001; Masliah and Milgram, 2000; Ruddle and Jones, 2001a; Wang and MacKenzie, 1999; Ware and Rose, 1998; and Zhai and Milgram, 1998].

Most CVE research has focused on the technical problems of delivering scalable environments to geographically disparate places, leading to the development of systems such as DIVE, MASSIVE and NPSNET [Frécon and Stenius 1998; Greenhalgh and Benford 1995; Macedonia et al. 1994], or the building of prototype applications [e.g., Benford et al. 2000; Sonnenwald et al. 2001]. Recently, a substantial amount of attention has turned to investigations of users' behavior when they interact together within CVEs, using tasks that ranged from moving a ring along a wire, to designing a room layout, solving puzzles, or carrying a stretcher [Basdogan et al., 2000; Hindmarsh et al., 2000; Sallnäs et al., 2000; Slater et al., 2000, 2001; Wideström et al., 2000]. However, with the notable exception of Basdogan et al., Sallnäs et al., and Slater et al [2001] there is a distinct lack of studies in which multiple users have had to cooperate to manipulate the same object, rather than communicate and manipulate different objects in a CVE.

The overall motivation for the present study is to investigate in detail the behavior of people when they perform a straightforward, shared, practical real-world task, in a CVE. This is a topic about which little is currently known, and the data from the study have wide-reaching importance in extending our understanding of the extent to which people can collaborate within VEs, and their behavior when they do so. In terms of application, studies of cooperative manipulation have particular relevance to simulation and training, and design reviews and data exploration. Within simulation and training, VE systems are used to mimic certain aspects of real-world operations. While current technology places limitations on the fidelity with which users can interact (e.g., the lack of locomotion devices and extended-range haptic feedback), astronauts can be trained in procedures for extravehicular activity even when not co-located [Loftin, 1997] and manufacturing designers could gain insights into the ergonomic problems of a design by "being" virtual humans and simulating together operations of manual materials handling (MMH) such as the installation of a dashboard into an automobile or the evacuation of a casualty [see Hubbard & Keates, 2000]. The role of collaboration in design reviews (e.g., in manufacturing) and data exploration (e.g., in the oil and gas industry) is one of promoting interplay and the exchange of ideas between pairs or small groups of people. Here, by moving the process of interaction from being one-sided ("I do this, while you watch") to

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

being truly cooperative, there is great potential for speeding up communication, ideas testing and information discovery.

The particular focus of the present experimental study is on cooperative object manipulation in VEs that are cluttered. That is, VEs which contain obstacles that impede a person moving through the VE and restrict the way in which objects can be manipulated. The task chosen for the study is known as the *piano movers' problem* [e.g., see Lengyel, Reichert, Donald, & Greenberg, 1990]. This is the generic task of maneuvering a large object through a restricted space, for example, part of a building. It is an ideal candidate for studying cooperation between multiple users in a CVE because the task is familiar to most people and can be varied in difficulty by changing the size or shape of the object being carried, or the layout of environment that it is being carried through. A typical scenario involves two or more virtual humans (3D mannequins) all carrying the same virtual object. Each user controls the position and orientation of "their" virtual human and its hands. Manipulations of the object are calculated by integrating together the movements of each virtual human and their hands, according to some pre-defined *rules of interaction*. The next section outlines the background to the study in more detail, paying particular attention to the rules that can be used to integrate different users' actions.

2. COOPERATIVE MANIPULATION

A framework put forward by Margery et al. [1999] identifies three levels of cooperation. In the first, users co-exist in a CVE, and can perceive and communicate with each other. In the second, each user can individually modify the contents of a scene, and in the third (Level 3) the users can simultaneously act on the same object. Within Level 3, a distinction is made between situations when users act on an object in an independent manner and the actions of each user have no direct effect on the others (e.g., one user modifying the object's position and another its color), and when the users' actions are codependent. The latter is the most pure form of cooperative manipulation, and is the case when two users work together to perform tasks such as moving a ring along a wire, the piano movers' problem, carrying materials around a virtual factory, or simulate the removal of a casualty on a stretcher.

Research into Level 3 cooperation is essential to inform our understanding of the extent to which people can cooperate to solve problems in CVEs, and has a direct impact on many different types of VE application (see above). To comprehensively study any type of interactive task, a wide variety of factors need to be investigated, including the

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

graphical techniques used to render the visual scene (texture mapping maybe considered the norm, but are shadows also present?), characteristics of the display such as its type (e.g., desktop, head-mounted, or CAVE) and field of view, the physical devices that are used for interaction (e.g., mouse vs. 3D prop [Hand 1997]), and the algorithms and parameters that are used within the interface software (e.g., zero- vs. first-order control, discrete vs. continuous activation, and the control gain). Clearly, the scope of any one study is limited to a small number of these factors and, although many have equal relevance to both single-user and cooperative manipulation, some are primarily applicable to cooperative manipulation. The cooperative factors fall into three categories: (a) those associated with network communications (these exist for all CVEs but are probably most severe for Level 3 cooperation), (b) feedback about the actions of each user that helps explain the resultant behavior of the object that is being manipulated, and (c) the ways in which multiple users' actions are integrated. Changes to each of these will modify users' behavior, and the speed and efficiency with which they can perform manipulation tasks.

2.1 Network communications

The key problems associated with network communications are centered round lag. As well as interrupting the flow of interaction, it causes technical difficulties in the implementation of a CVE system. One difficulty is that the CVE has to determine whether multiple interaction requests that are received within a short space of time should be considered to be simultaneous (i.e., synchronized) or successive [Broll 1995]. A second is dealing with packets of data that are lost on the network, which can cause severe inconsistencies between different local representations of a CVE. For example, if the packet containing a "stop moving " event is lost then a user may be stationary in one copy of CVE but moving forwards in another [see Slater et al., 2001]. A third difficulty relates specifically to haptic feedback. Haptic feedback typically uses an update rate of 1 kHz but this is not achievable in distributed systems. In a study that used an Internet2 link between the UK and the USA, a haptic update rate of 60 Hz was achieved but this led to a simple cooperative task (lifting an object off a table) falling apart after just a few seconds, because the users' two haptic scene graphs fell out of synchronization [Slater et al., 2001].

In common with some previous studies [e.g., Basdogan et al., 2000; Hindmarsh et al., 2000; Sallnäs et al., 2000], we circumvented most of the problems associated with lag by running our CVE on a single host computer. This allowed us to focus on the behavioral aspects of interaction, using a "best-case scenario" for network communications.

2.2 Feedback

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

Feedback is essential in cooperative manipulation if users are to understand the intended actions of each other. It has been shown that haptic feedback significantly improves participants' performance during cooperative manipulation [Basdogan et al. 2000; Sallnäs et al., 2000] but, unfortunately, there are major barriers to its use. First, network delays currently preclude the use of haptic feedback in even simple manipulation tasks that use distributed systems, because of the technical difficulty of maintaining synchronization between multiple copies of haptic scene graphs (see above). In the studies by Basdogan et al. and Sallnäs et al. this was overcome by using a single host computer and having participants located nearby each other. Second, current haptic devices such as the PHANTOM have a small working volume and only the very latest versions can provide force feedback for rotational DOFs. In addition, this rotational feedback cannot be provided over the full 360 degrees of rotation about any particular axis. One solution to these limitations is to scale movements of the physical interface (e.g., translations of the Phantom) so that they produce correspondingly larger movements of a virtual object, but this introduces haptic instabilities. A second solution is to provide a clutch that allows the user to reposition and reorient the haptic device relative to the virtual object, and informal tests using a single-user VE indicate that this approach has considerable promise [McNeely et al., 1999; personal communication, W. A. McNeely, 27 November 2001].

Alternatives to haptic feedback come from providing visual or auditory feedback. Visual feedback can provide a variety of types of information such as indicating the position to which each user is attempting to manipulate an object, the velocity of their hand movements, or the forces that they attempt to apply. Different visual representations of each of these could be provided, for example, a wireline outline showing the attempted position of the whole object, or just the part of the object nearest the user's grasp point on the object. Alternatively, a semi-transparent version of the object could be rendered to provide similar information. Auditory feedback can be used to provide information about discrepancies in the forces or movements that users attempt to apply, and the most useful types of technique are likely to use sound volume or pitch to indicate the magnitude of any discrepancy. Whether visual or audio techniques are used, the most effective form of feedback is best chosen using pilot studies for particular applications and types of environment.

2.3 Action integration

It is extremely difficult for multiple users to manipulate a single virtual object in exactly the same way because even if they are carrying the object "together", they are not

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

physically connected by the object. A similar problem arises if an individual user is manipulating a virtual object using two hands. In both cases, rules have to be defined that allow multiple inputs to be integrated. The primary focus of the present study is on these rules of interaction, which represent one of the most fundamental aspects of user interfaces for cooperative manipulation. First, however, we put these rules in context by considering real-world interaction.

The two forms of real-world interaction that we consider are bimanual (one person using both hands) and cooperative (multiple users). Bimanual interaction can take place on either a symmetric or an asymmetric basis. With symmetric interaction both hands perform the same role, for example, lifting an object or turning it through a large angle. In asymmetric interaction the hands perform essentially different roles and one example is when the non-dominant hand provides a frame of reference within which fine manipulations are performed by the dominant hand [see Guiard, 1987]. In practice, people switch between symmetric and asymmetric control with little conscious thought, and the process of switching is supported by subtle changes in a person's grasp.

Object manipulation by multiple people shares many similarities with bimanual interaction because, again, manipulation may take place on either a symmetric or an asymmetric basis. Symmetric manipulation requires two (or more) people to perform actions that are coordinated in all respects (i.e., actions that have the same magnitude and direction as each other, and are performed at the same time), but this hides some of the subtlety of physical human-human interaction. Consider two people who are carrying a long pole, each holding one end. If one person holds the pole in a fixed position then it cannot be moved by the other, provided that the first person is strong enough. Although the two people make similar movements (neither ends up moving their hands), the movements that they intended to make (hold the pole still vs. move it) are completely different. In another situation, the movements of one person could be guided by those of the other, via forces transmitted through the pole. However, with asymmetric manipulation two users will deliberately make substantially different actions, for example, one person allowing an object to pivot while the other changes its orientation and, as with bimanual interaction, switching between the two forms of interaction is aided by the ease with which people can change their grasp.

The rules that can be used to integrate multiple users actions in VEs (or bimanual interaction by a single user) fall between two extremes. At one, an object can only be moved if the users manipulate it in exactly the same way, whereas at the other the object

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

moves according to some aggregate of all the users manipulations. The extremes correspond to symmetric and asymmetric manipulation, respectively.

In practice, users can never manipulate a virtual object in exactly the same way as each other and this leads to two main options for symmetric action integration. One allows manipulation to take place provided the users' actions are within a certain, small tolerance of each other, whereas the other option simply uses the common component of the users' actions and is akin to the dot product of vector mathematics. It is this latter option that has been implemented in most CVE interfaces to date. One example was the study by Basdogan et al. [2000], in which pairs of participants cooperated to move a virtual ring along a virtual wire and haptic feedback was provided using two PHANToMs (SensAble Technologies Inc.). The ring moved by an amount that was proportional to the common component of the forces that each participant applied via their PHANToM, with any residual forces (the non-common component) being ignored by the VE software. This made the task substantially easier than it would have been if imbalances in the forces applied by the two participants had caused the ring to rotate. In another study, users lifted up virtual cubes by pushing on opposite faces using PHANToMs and, again, any torques that would have been produced by pushing at different places on the faces were ignored by the VE software [Sallnäs et al., 2000].

Asymmetric manipulation allows each user to manipulate an object in a different way, subject to the constraint of both users maintaining their hold on the object. One implementation of this was described by Fröhlich et al. [2000] who developed an algorithm for bimanual interaction that can also be applied to multiple users. Relative movements of a user's hands were used to calculate imaginary forces (no haptic feedback was provided) and these were then used to define the manipulations made to an object. If only one hand was moved then the object moved by an amount that was approximately half that made by the hand, and if the user then moved their other hand by the same amount the object moved the remainder of the way. Variations of asymmetric action integration include those where multiple users are not treated as "equal partners."

3. RESEARCH QUESTIONS

To date, interfaces for cooperative manipulation have always been implemented so that they allow either symmetric or asymmetric action integration, but not both. In addition, no studies have been made that compare these forms of integration, so it is not clear whether any benefit would be gained by providing both. The primary purpose of the

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

present study was to compare participants' performance when their actions were integrated either symmetrically or asymmetrically. The following hypotheses were made:

- (H1) Overall, participants would perform the task quickest when asymmetric action integration was used.

- (H2) Asymmetric integration would prove superior to symmetric integration even for the components of the manipulation task for which participants needed to make similar actions (e.g., rotating an object in a doorway). This was predicted because symmetric integration imposes the need for users to synchronize their actions, whereas asymmetric integration is more flexible and allows users to adapt their actions to suit the nature of the task.

There were two additional aspects of the study, namely: (a) a comparison of the time that participants took to perform the task cooperatively with the time taken in a previous study when participants performed the same task individually, and (b) detailed analysis of participants' actions during interaction, focusing on the extent to which their actions were coordinated.

4. EXPERIMENT

The experiment used a cooperative object manipulation task to study the effects of interface algorithms that allowed either symmetric or asymmetric interaction on participants' performance. The task that was used is known as the piano movers' problem and involved pairs of participants in moving a large virtual object through two VEs that contained parts of a virtual building. Participants made their interactions via a pair of virtual humans that were situated in the VE. Two different types of interface algorithm were compared, which represent the two extremes of action integration that were outlined above. One algorithm allowed only the *synchronized* component of participants' manipulations to take place (symmetric interaction), and the other allowed the *mean* (asymmetric). One of the VEs contained two openings that were *offset* from each other, and the other was a *C-shaped* section of corridor. In some parts of these VEs the participants had to make similar movements to each other, but at other times the movements were substantially different (e.g., one participant maneuvering themselves through a narrow opening while the other traveled down a wide corridor). A repeated measures design was used, with each pair of participants at different times performing the

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

task using both rules of integration. Participants' interactions with the VE system were recorded in real time, allowing both their performance and their behavior to be analyzed.

4.1 Method

4.1.1 Participants. Twenty participants (11 men and 9 women) took part in the experiment. Their mean age was 22.6 ($SD = 4.5$). All the participants volunteered for the experiment, were paid an honorarium for their participation, and had previously successfully completed another experiment in which they performed similar tasks, but in a single-user mode [Ruddle, Savage et al., in press]. Participants performed the experiment in pairs, with some pairs being male only, others female only, and the remainder mixed. The pairs were divided into two groups to counterbalance the order in which the two rules of interaction (synchronized or mean) were used.

4.1.2 VE Application. The VE software was a C++ Performer application that was designed and programmed by the authors, and ran on an SGI Maximum IMPACT workstation. This drove the display for both participants in each pair, using the Impact Channel Option to divide the graphics frame buffer into two VGA outputs (the view for each participant) that were supplied to two 86 cm (34-inch) monitors. The application update rate was 15 Hz.

The layout of the laboratory used for the experiment is shown in Figure 1. The participants in each pair stood back-to-back, facing a monitor, and separated by a wooden partition. Participants were allowed to talk to each other but could not see each other. All they could see was a view of the VE, which showed the walls and floor of the environment, and two virtual humans that were carrying a large object. Interior and plan views of the two environments that were used are shown in Figures 2, 3 and 4. For details of illustrative videos, see Appendix A.

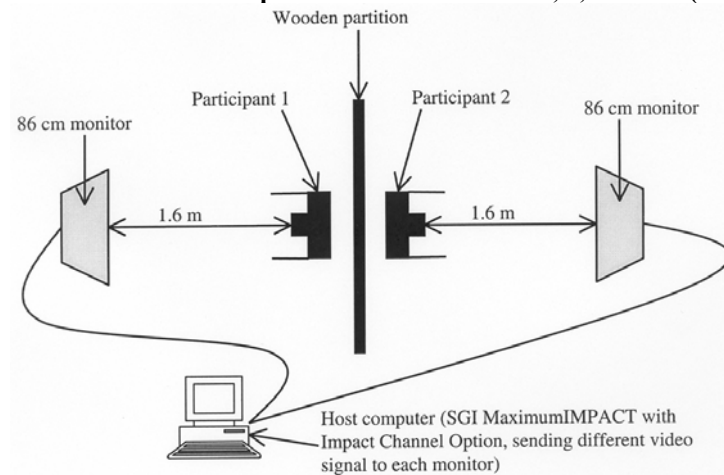


Fig 1. Layout of the laboratory used in the experiment.

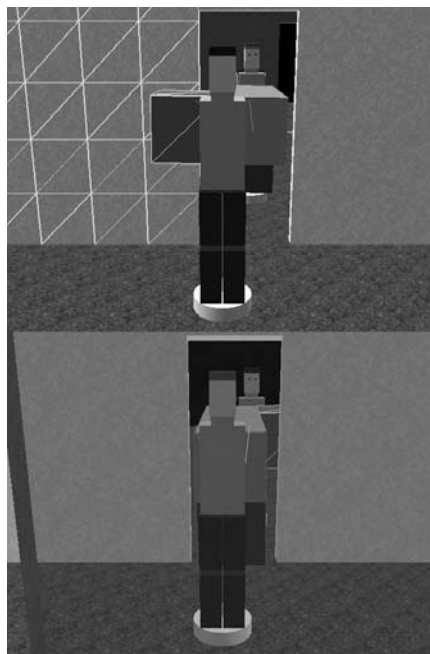


Fig 2. A view inside the offset VE showing the view seen by Participant 1 (top) and Participant 2 (bottom) of the same setting. The wireline highlighting indicates that a collision is taking place between the wall and the stub of the object held by Participant 1.

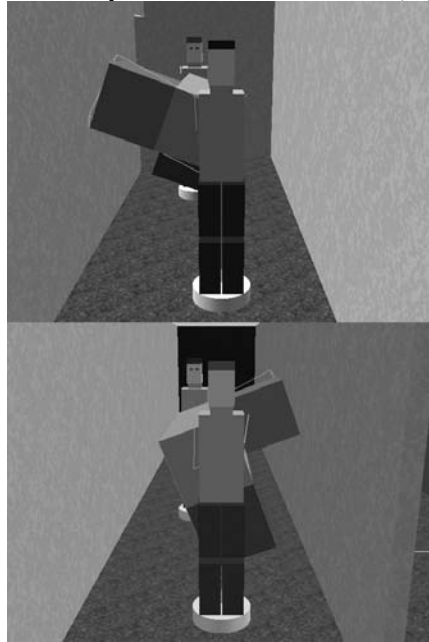


Fig 3. A view inside the C-shaped VE showing the view seen by Participant 1 (top) and Participant 2 (bottom) of the same setting. The wireline image of the end of the object shows where Participant 1 is trying to manipulate the object to (synchronized rule of interaction).

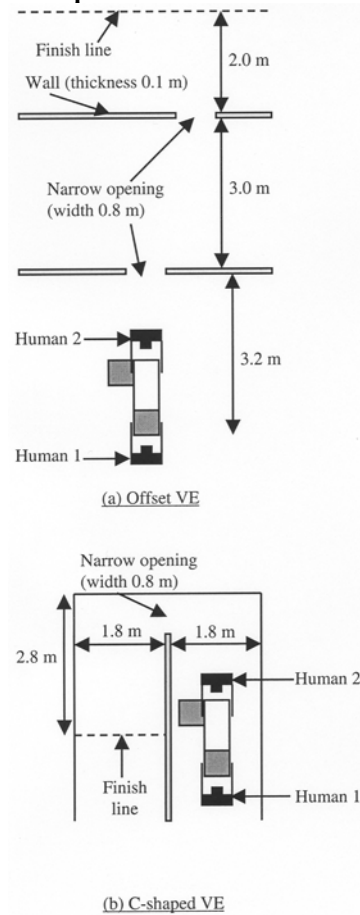


Fig 4. Plan views of the offset (a) and C-shaped VEs (b). In both cases, the ceiling was at a height of 2.4 m and the narrow openings were 2.0 m high. Human 1 moved forwards and was controlled by Participant 1. Human 2 moved backwards and was controlled by Participant 2.

The two virtual humans were the virtual counterparts of the two participants (their embodiments within the VE). Each participant's viewpoint was positioned 3 m behind the position of their virtual human, connected by an egocentric tether (an "over the shoulder" view). This meant that the participant's direction of view was always the same as that of the virtual human but the participant was able to see the human's immediate surroundings in the VE, despite the impoverished field of view ($48^\circ \times 36^\circ$). This type of view perspective is under investigation for displays of aircraft navigation information [Wickens & Prevedt, 1995] and has been used with great success in a number of earlier VE studies [e.g., Hindmarsh et al., 2000]. Had a human's-eye view been adopted instead then a participant would not have been able to see the whole of the object and the other participant's virtual human in a single view. Because of the offset of the tether, a

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

participant's viewpoint was sometimes on the opposite side of a wall to their virtual human. When this occurred, the walls in question were rendered as semi-transparent using an alpha value of 0.2 (0.0 and 1.0 were fully transparent and opaque, respectively).

The object was an abstract shape, similar to one of those used in the earliest studies of mental rotation [Shepard and Metzler 1971]. Its dimensions are shown in Figure 5.

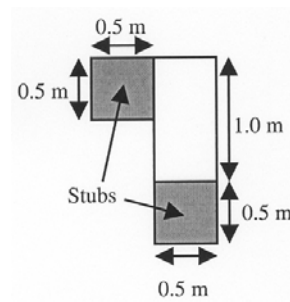


Fig 5. Dimensions of the Shepard-Metzler object used in the experiment. The stubs were the same size as, but at 90 deg. to, each other.

4.1.3 User Interface. A variety of design decisions have to be made when any type of user interface is implemented for a VE. In general, the aim is for user interaction to be as efficient as possible, and this typically occurs when interaction takes place in a "natural" manner. For example, it is well known that there should be correspondence between users' physical and virtual hand movements, as was implemented for the present experiment, subject to the limitations imposed by factors such as collisions and action integration. Ideally, users should also be able to travel through a VE by walking and physically turning around. Walking interfaces remain under development (see Hollerbach, 2002), but physical rotation is straightforward if a head-mounted display (HMD) is used. Unfortunately, HMDs bring with them the problem of VE sickness, which is particularly problematic when VEs are used for long periods of time. It was primarily for this reason that monitor displays were chosen for the present experiment, although it should be noted that a single-user pilot study performed using a monitor and an HMD indicated little in the way of performance differences once participants were fully trained. Other aspects of VE interface design such as the implementation of a clutch are not natural, but are practical ways of overcoming the fixed position of buttons on interface devices, and are known to aid user performance.

In the present experiment, each participant held an interface prop, the position and orientation of which was tracked using a Polhemus Fastrak sensor and the MR Toolkit

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

[Green 1995]. The props were small boxes (100 x 75 x 40 mm) that had the sensor mounted on the top, and four buttons (two on the top and two on the front). If a participant held down one button they accelerated forwards (i.e., in their direction of view) at 0.5 ms^{-2} , to a maximum speed of 0.5 ms^{-1} , and if they held down another button they accelerated backwards at the same rate. The third button acted as a clutch that allowed participants to reposition and reorient the prop and, therefore, their hands without changing the position or orientation of the object. The fourth button was used to change the mode of the Fastrak sensor. When the button was held down, changes of the prop's orientation caused the participant's direction of view to be rotated. If the third and fourth buttons were held down simultaneously then the participant's virtual hand position remained fixed but their body was repositioned according to their physical hand movements. This allowed participants to move their virtual humans directly in any direction and was particularly useful in the offset VE because it allowed the virtual humans to sidestep between the two openings.

Throughout the duration of each trial, the two virtual humans grasped the two ends of the object. The details of manipulation are described first for single-user interaction, followed by the modifications made for the two types of twin-user interaction.

In general, there was 1:1 correspondence between the physical movements of a participant's hand and the movements that their virtual counterpart attempted to make. Exceptions to this occurred when the participant was using the clutch, sidestepping, or there was a collision. When the clutch was used the object and the virtual human both remained stationary while the participant physically changed the position and orientation of their hands. Within the VE software, this discrepancy was accounted for by position and orientation offsets, but subsequent manipulations of the object still took place as if the participant's hands were in the same place as those of their virtual human (hand-centered manipulation [Bowman and Hodges, 1997]). Sidestepping worked in the same way.

Collisions of the object were detected using the RAPID software library [Gottschalk et al. 1996]. If the object collided with the environment then it was prevented from moving (it was not allowed to penetrate the walls, floor or ceiling) but a participant could still reposition their virtual human relative to the object (a stop-by-parts collision response algorithm; see below). Graphical highlighting indicated which geometric (e.g., tri-strip) primitives were in collision and the offsets between participant's physical and virtual hand position and orientation were redefined. If the object collided with a virtual human then both the object and that human were prevented from moving. As with object-environment

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

collisions, graphical highlighting indicated which primitives were in collision and the physical-virtual hand offsets were redefined. If a virtual human collided with the environment then a "slip" response algorithm [see Ruddle & Jones, 2001b] allowed the virtual human to continue moving tangentially to the colliding surface. The choice of different response algorithms for the object and virtual humans represented the fact that, in real life, it is trivial for people to avoid walking into walls, but if an object is scraped against a wall then damage tends to occur.

Collisions of the object with the environment could be handled by two types of response algorithm: *stop-by-parts* or *stop-as-a-whole*. Stop-as-a-whole is the type of algorithm that is most often implemented in VEs, and means that the position of all objects in a scene is temporarily frozen if a collision takes place anywhere in a particular graphics frame. Stop-by-parts only freezes the position of the objects that are actually in collision. This type of algorithm is a substantially more complex to implement, but greatly increases the ease with which users can interact. In our earlier studies of single-user object manipulation in cluttered VEs, participants performed piano movers' trials in the C-shaped VE 33% quicker when they used stop-by-parts than when they used stop-as-a-whole [Ruddle, Savage et al., in press]. In twin-user manipulation stop-by-parts would have an even greater advantage because each user can vary the position and orientation of their virtual human, irrespective of whether the other user's virtual human is in collision, or the object is in collision. For these reasons, stop-by-parts was used throughout the present study.

With twin-user interaction, the process of movement was broken down into two stages. First, the translational movements of the two virtual humans were considered. The raw movements of the two humans were combined by taking either the mean or the dot product (mean and synchronized rules, respectively), and this ensured that the two humans did not drift apart even though their speed or direction of movement usually differed. With mean movement, each participant could move both humans and the object through the VE, even if the other participant's human was attempting to remain stationary. However, with synchronized movement, progress would only be made if both participants moved their respective humans in non-opposing directions. The second stage of movement was object manipulation. With this, as with the humans' movements, the resultant movement was either the mean or dot product of the raw hand movements made by the two participants, as measured by the sensors on the interface props. With mean movement, each participant could manipulate the object by themselves, but with

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

synchronized movement the participants had to coordinate their manipulations in real time. Translational movements of the virtual humans and the objects were calculated using the same algorithms, details of which are shown in Figure 6. Changes of the object's orientation were calculated using similar algorithms.

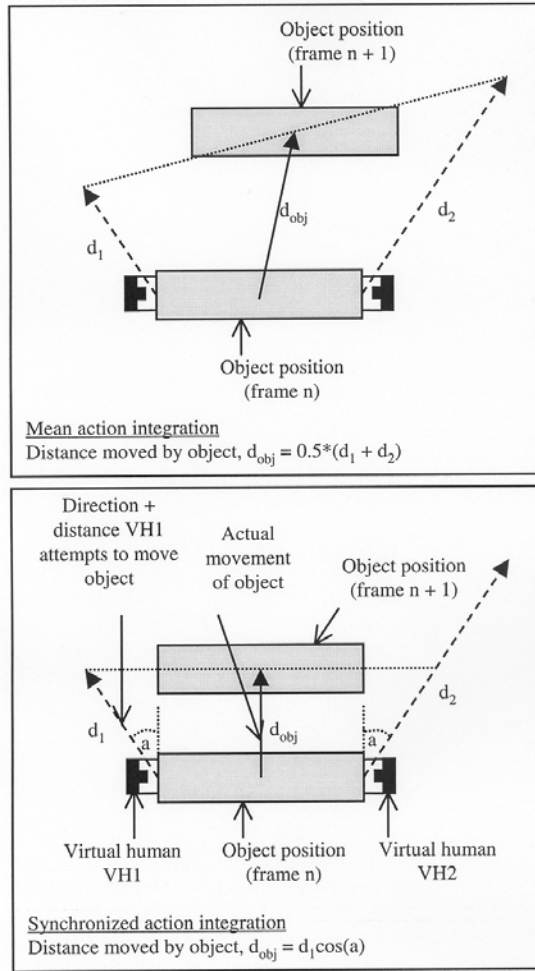


Fig 6. Effect of synchronized (bottom) and mean (top) rules of interaction on the manipulation of an object when two users attempt to move it in differing directions and by different amounts. In both cases, sensor readings allow the change in participants' raw hand position to be calculated (the vectors d_1 and d_2). For mean interaction the object's movement is the mean of d_1 and d_2 . For synchronized interaction the object's direction of movement bisects the angle made between the manipulations of each individual user, and the distance moved is calculated from the dot product. The mean and synchronized rules would only produce identical movement of the object if the users' actions were completely synchronized in time, direction, and magnitude.

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

Clearly, there was usually a discrepancy between the manipulations that each participant attempted to make to the object, and those that actually took place. Graphical feedback, in the form of a wireline model of each participant's end of the object, indicated participants' attempted manipulations (see Figure 3).

4.1.4 Procedure. Participants were run in pairs and performed the experiment over two separate days. On the first day they performed trials in the offset VE and on the second they performed trials in the C-shaped VE. Over the two days participants took an average of six hours to perform the experiment, including rest periods to help alleviate fatigue.

At the start of the first day, the experimenter demonstrated how to perform the piano mover's task, using a physical scale model of the object and the offset environment. Then the participants practiced moving the object through the offset VE in single-user mode. For this, the experimenter demonstrated how to move the object and then each participant performed two practice trials. In each trial, a participant carried the object from the starting position until both virtual humans had crossed the finishing line, which was marked on the floor of each VE (see Figure 4). It is important to emphasize that all of the participants were already familiar with the experimental task because they had previously taken part in an experiment that studied single-user interaction for the piano mover's problem. That is, the practice trials acted as a reminder rather than training *de novo*.

After the single-user practice, the participants performed trials in twin-user mode. First they performed three practice trials using one of the interfaces (e.g., synchronized), and then three practice trials using the other interface (e.g., mean). Then they performed six test trials using the first interface, and then six test trials with the second interface. Each set of test trials was split into two blocks of three.

The format of the second day was identical to the first, except the C-shaped VE was used. Throughout all of the twin-user trials, on both days, each participant used the same virtual human. One of these moved forwards to carry the object, and the other moved backwards. During the practice trials, the experimenter gave advice on how to perform the task, but during the test trials the experimenter was silent. If participants had not completed a test trial after 600 s then the trial was terminated and they progressed to the next trial.

4.2 Results

Three sets of analyses are reported. The first concerns the time that participants took to complete the test trials. The second compares the time data from the present study with

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

data from a similar study in which participants performed the same task but in a single-user mode. The third refers to participants' behavior during the trials and the extent to which they coordinated their interactions. One pair of participants timed out in two test trials that used the synchronized rule in the C-shaped VE, and two pairs of participants timed out in one trial each when using the mean rule in the C-shaped VE. For the analyses below, the times for these trials were set to 600 s. Statistical comparisons between the offset and C-shaped VEs are not reported because the task was designed to be more difficult in the latter. None of the interactions was significant.

4.2.1 Time Data. First, participants' learning was investigated by analyzing the time taken to complete each of the test trials. An analysis of variance (ANOVA) that treated the interaction rule (mean or synchronized), environment (offset or C-shaped), and trial number as repeated measures showed that participants took less time as the trials progressed, $F(5, 45) = 3.73, p < .01$, but there was no overall difference between the two rules, $F(1, 9) = 0.64, p = .44$ (see Figure 7).

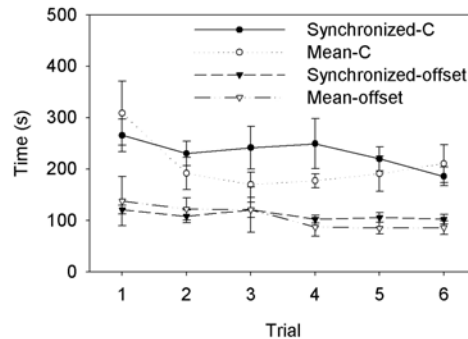


Fig 7. Mean time taken to complete the test trials in the four combinations of interaction rule and VE. Error bars indicate the standard error (SE).

The remainder of the analyses involving the time data used participants' mean performance in the second block of test trials (Trials 4 - 6) because this discounted the effects of learning that were most marked in the early test trials. Unless otherwise stated, the data were analyzed using repeated measures ANOVAs that treated the rule and environment as repeated measures. The mean time that participants took to complete Trials 4 - 6 was similar with two rules, $F(1, 9) = 1.14, p = .31$. However, to allow a more detailed analysis of these times, each trial was broken down into stages when: (i) the object was in collision with the structure of the environment (the walls, floor and ceiling)

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

or either virtual human, (ii) both virtual humans were stationary (i.e., a resultant speed of zero; no translational movement of their bodies) but the object was not in collision, or (iii) the humans were moving. In (i) it is important to note that the humans' resultant speed was set to zero whenever the object was in collision. Also, with the synchronized rule, the humans only moved if both participants attempted to move, but with the mean rule both humans would move if either participant attempted to move. Repeated measures ANOVAs showed that the object was in collision for more time with the mean rule than with the synchronized rule, $F(1, 9) = 10.47, p = .01$, but the humans were stationary with the object not colliding for less time with the mean rule, $F(1, 9) = 23.30, p < .01$. The difference in the time for which the humans were moving was not significant, $F(1, 9) = 0.85, p = .38$ (see Figure 8).

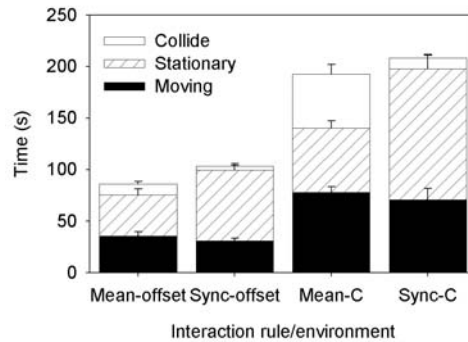


Fig 8. Mean time spent with the object colliding, both virtual humans stationary, or the humans moving through the VE in the second block of test trials. Sync = synchronized rule. Error bars indicate the SE.

Participants took twice as long to complete the trials in the C-shaped VE as in the offset VE. To provide information on where participants experienced difficulties in the former environment, each trial was broken down into five phases. These were when: (1) both virtual humans were traveling towards the narrow opening, (2) Human 2's end of the object was being maneuvered through the opening, (3) the object was being rotated with one human on either side of the opening, (4) Human 1's end of the object was being maneuvered through the opening, and (5) both humans were traveling towards the finish line. Human 1 was the virtual human that moved forwards while carrying the object, and Human 2 was the virtual human that moved backwards. Both of these, together with the position of the narrow opening are shown in Figure 4b. The percentage of time in each trial that participants spent in each phase was calculated and analyzed using ANOVAs

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

that treated the interaction rule as a repeated measure. Participants performed Phase 3 significantly quicker with the synchronized rule than the mean rule, $F(1, 9) = 5.92, p = .04$. There were also marginal effects for Phase 1, $F(1, 9) = 4.56, p = .06$, and Phase 2, $F(1, 9) = 3.73, p = .08$, indicating in both cases that participants were slower with the synchronized rule than the mean rule. The differences for Phases 4 and 5 were not significant (see Table I).

Table I. Mean (*SD*) percentage of time that the object was in each part of the C-shaped VE for the second block of test trials. The positions of Human 1 and 2 are shown in Figure 4b.

Trial phase	Rule		$F(1, 9)$	p
	Mean	Synchronized		
1. Traveling towards opening	23.9 (6.7)	29.0 (6.2)	4.56	.06
2. Human 2's end maneuvered through opening	18.5 (7.2)	22.5 (5.9)	3.73	.08
3. Rotating object in the opening	24.5 (11.7)	14.3 (2.6)	5.92	.04
4. Human 1's end maneuvered through opening	16.8 (5.8)	16.1 (7.7)	0.22	.65
5. Traveling towards finish line	16.4 (3.0)	18.1 (7.0)	0.54	.48

4.2.2 Single- vs. Twin-user Interaction. Participants were anticipated to take longer to complete the trials than they would have if they had performed the task in single-user mode. To provide information on the magnitude of this difference, the data from the present study were compared with the data from an earlier study [Ruddle, Savage et al. in press]. That study used the piano movers' problem to investigate the effect that different rules of interaction, for example, stop-by-parts vs. stop-as-whole collision response, had on object manipulation in restricted virtual spaces. The results from the single-user study were used to optimize interaction in the twin-user study, for example, by the implementation of a twin-user version of stop-by-parts (see Section 3.1.3 User Interface). The analyses reported below compare participants in the present study with the data for the 15 participants who were in the best performing group of the single-user study.

Separate ANOVAs of mean trial time in the second block of test trials were performed to compare each twin-user rule with the single-user data for each environment. The mode

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

(single-user vs. synchronized (or mean) interaction) was treated as a between participants factor. The mean-offset twin-user trials took a similar amount of time as the single-user trials, $F(1, 23) = 0.13, p = .72$. However, the twin-user task took significantly longer than the single-user task for the synchronized-offset, $F(1, 23) = 4.41, p = .05$, mean-C, $F(1, 23) = 5.53, p = .03$, and synchronized-C trials, $F(1, 23) = 7.91, p = .01$ (see Figure 9).

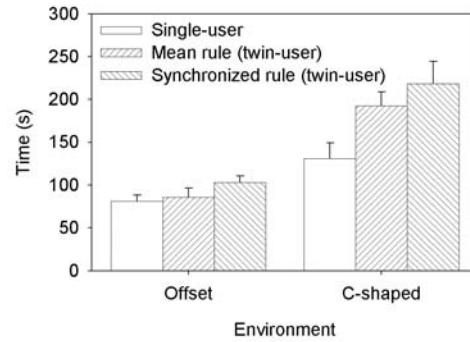


Fig 9. Mean time spent completing the trials in single- and twin user mode. The single-user study is reported in Ruddle, Savage et al. [in press]. Error bars indicate the SE.

4.2.3 Interaction Behavior. In any form of cooperative (virtual) object manipulation there is a difference between the movements that each participant makes to the object and those made by their virtual counterpart, because the relative movements of participants are not constrained by the "rigid" object they are holding. This difference depends on the extent to which the participants' movements are coordinated. The data reported below provide information on this coordination and refer to the second block of test trials.

4.2.3.1 Clutch and Sidestepping. The first set of behavioral data refer to the amount of time that participants spent either using the clutch to physically reposition their hands without manipulating the object, or sidestepping, which also precluded manipulation of the object. For the majority of each trial, both participants were manipulating the object in "normal" mode (no clutch or sidestepping). There were also substantial amounts of time when one or other participant was using the clutch or sidestepping, but participants performed these actions together much less frequently (see Figure 10).

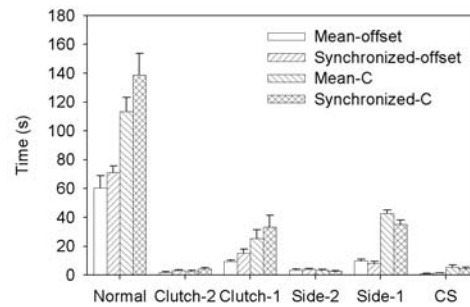


Fig 10. Mean time spent in each type of manipulation mode in the second block of test trials. Clutch-2 = both participants using clutch. Clutch-1 = one participant using clutch and the other in normal mode. Side-1/2 = one/both participants moving sideways. CS = one participant using the clutch and the other moving sideways. Error bars indicate the SE.

4.2.3.2 *Virtual Human Movement Coordination.* The next data relate to the extent to which participants coordinated the raw translational movements of their virtual humans. Each time step (a graphics frame) in the trials was classified according to whether: (i) participants were attempting to move their humans in the same direction (i.e., one forwards and the other backwards in their respective body spaces), (ii) participants were attempting to move their humans in opposite directions, (iii) neither participant was attempting to move their human, or (iv) one participant was attempting to move their human but the other participant was attempting to keep their human stationary. These data are illustrated in Figure 11.

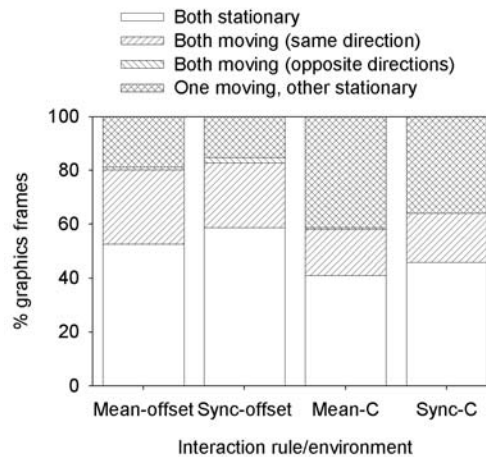


Fig 11. Percentage of the graphics frames in which participants were attempting to keep their virtual human stationary, or to move it (second block of test trials). Sync = synchronized rule. Error bars indicate the SE.

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

Of particular interest is the amount of time for which one participant was attempting to move their human while the other participant was not. In interpreting these data it should be noted that the resultant movement was different for the two rules of interaction. With synchronized interaction neither human would move if either one attempted to remain stationary, but with mean interaction both would move if one attempted to. In other words, the combination of stationary and moving inputs can be considered as a "valid" form of interaction with the mean rule. It is also probable that use of the mean rule also influenced participants' behavior when they used the synchronized rule. However, this stationary-moving behavior occurred frequently both with participants who performed the test trials first using synchronized interaction, and with those who used the two rules of interaction in the opposite order.

4.2.3.3 Hand movement Coordination. The final sets of data refer to the coordination of participants' hand movements. Translational and rotational movements were analyzed separately, but in a similar manner. First, graphics frames in which either participant was using the clutch or sidestepping were discarded. Next, the speed at which participants moved their hands in a frame was calculated, irrespective of the direction of movement, and frames in which neither participant moved quicker than 0.1 ms⁻¹ (translation) or 20 degrees/s (rotation) were discarded. This ensured that only major movements were considered and sensor noise could be ignored. Overall, 23 % and 35 % of frames remained for the translation and rotation analyses, respectively. For these, the speed of the slower user was expressed as a percentage of the speed of the faster user, and averaged for all of the non-discarded frames in a trial. An ANOVA that treated the rule (mean vs. synchronized) and environment (offset vs. C-shaped) as repeated measures showed that the speed of participants' translational hand movements was more coordinated with the synchronized rule than with the mean rule, $F(1, 9) = 13.17, p < .01$, and more coordinated in the offset VE, $F(1, 9) = 12.93, p < .01$. Similarly, the speed of participants' hand rotations was more coordinated with the synchronized rule, $F(1, 9) = 10.55, p < .01$, and in the offset VE, $F(1, 9) = 13.93, p < .01$ (see Figure 12).

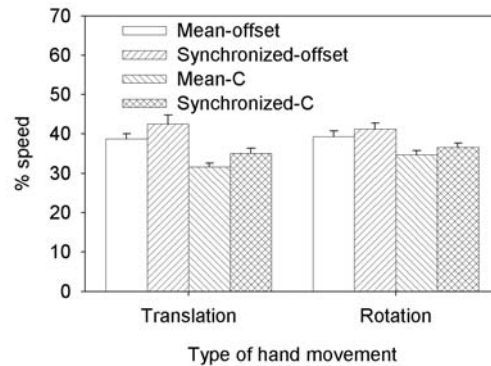


Fig 12. Mean coordination of the speed of participants' translational and rotational hand movements in each graphics frame in the second block of test trials. For each frame, % speed = 100 * speed of slower participant / speed of faster participant. Error bars indicate the SE.

The extent to which participants coordinated the direction of their translational and rotational movements was also analyzed. Frames in which either participant was using the clutch or sidestepping were discarded, as were frames in which either participant moved their hand at a speed of less than 0.05 ms⁻¹ (translation) or 10 degrees/s (rotation). For the remaining frames (12 % and 23 % for translation and rotation, respectively) the mean angle between participants' movements was calculated using the dot product for translations and the angle between the axes of rotations for rotational movements. An ANOVA showed a similar pattern of results to the speed coordination data. Participants' translational movements were more coordinated with the synchronized rule than with the mean rule, $F(1, 9) = 8.07, p = .02$, and more coordinated in the offset VE, $F(1, 9) = 25.77, p < .01$. For the hand rotations the difference between the two rules was not significant, $F(1, 9) = 2.18, p = .17$, but coordination was greater in the offset VE, $F(1, 9) = 45.54, p < .01$ (see Figure 13).

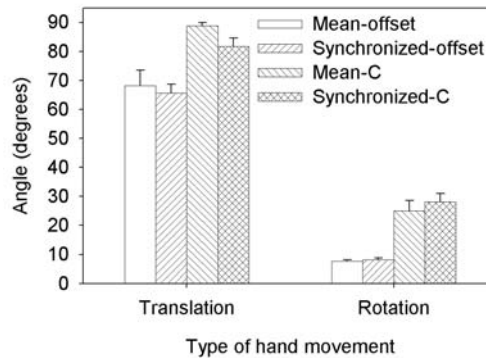


Fig 13. Mean angle between participants' translational and rotational hand movements in each graphics frame in the second block of test trials. Error bars indicate the SE.

5. CONCLUSIONS

The present study adopted a paradigm called the piano movers' problem to study two methods of integrating different users' actions during a cooperative object manipulation task. The basic problem was for pairs of participants to move a bulky virtual object through a restricted space. Participants' actions were integrated by calculating either the common (synchronized) component of their manipulations, or the mean. By using two different VE layouts (offset and C-shaped) the effect of task difficulty on cooperation behavior was also investigated.

Comparison of the present study's time data with the data for an equivalent single-user study shows the impact on participants of having to perform the task in cooperation with another user. This *cooperation overhead* was negligible in the offset VE when the mean rule was used to integrate participants' actions, but substantial in the C-shaped VE (47 % and 67 % for the mean and synchronized rules, respectively). It is to be expected that user performance deteriorates when a task such as the piano movers' problem has to be performed cooperatively, but this is the first time that such performance differences have been objectively measured for a CVE. These data also show that moderately difficult tasks can be performed cooperatively almost as quickly as when individuals perform them by themselves.

Contrary to expectations, and despite the opportunity for improvement provided by the magnitude of the collaboration overhead, there was no significant difference between the time that participants took to perform the task with the mean and synchronized rules. However, detailed analysis of the time data indicated that there were significant

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

differences for particular phases of the task in the C-shaped VE. These can be summarized as follows. If both participants needed to execute a similar action (Phase 3 of the trial; see Table 1) then manipulation was quickest if the interface only allowed synchronized movement to take place. However, if the participants needed to perform different types of movement (asymmetric interaction; e.g., one to maneuver their virtual counterpart through the door while the other virtual human just traveled down the corridor) then interaction was quickest with the mean rule.

Both of these differences can be explained. In Phase 3 with the mean rule, the manipulations made by each participant on their own may have led to the object being successfully rotated in the opening if the other participant had done nothing but, instead, they both tried to rotate the object and the slight differences in their manipulations caused the object to collide. With the synchronized rule, only the common, predominantly non-colliding component of rotation occurred. This indicates that a synchronized rule of interaction is beneficial for precision manipulation in a CVE. By contrast, it was difficult for participants to synchronize their manipulations of the object if they were simultaneously trying to perform different types of maneuver in the VE. In this situation performance was better with the mean rule. The advantage of the mean rule for asymmetric components of the task was predicted, but the advantage of the synchronized rule for other components was not and shows the advantages that can accrue to users if the VE system constrains the manner in which they can interact. This also illustrates the need for VE interfaces to allow flexibility in the rules chosen to integrate users' actions as tasks are varied.

The principal cause of the magnitude of the cooperation overhead and the difficulties that participants experienced with both forms of action integration was the lack of coordination between participants' movements. With the synchronized rule, one participant was attempting to move through the environment while the other was stationary, with the net result that neither moved, for up to 40% of the time taken to complete a trial (see Figure 11). With both rules, synchronization of hand movements was poor, particularly in the direction in which participants translated their hands (see Figure 13). With the mean rule, a direct consequence of the lack of coordination was the large proportion of time for which the object was in collision with the environment. Also of note was the fact that, although participants' hand movements were significantly more coordinated with the synchronized rule than with the mean, the magnitude of the difference between the two rules remained small. Thus, even though the object could only

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

be moved with the synchronized rule if participants did coordinate their actions, the level of coordination remained poor.

It is inconceivable that participants would have been uncoordinated to such a degree if it had been possible to provide haptic feedback. Thus, this study highlights the likely benefits to cooperative manipulation should haptic devices become capable of providing feedback over working volumes of the size needed to perform tasks such as the piano movers' problem. However, given that wide-range haptic devices will not become available in the immediate future, this study also identifies the need to develop ways of simulating haptic feedback for tasks that involve cooperative manipulation. The most likely candidates are visual or auditory information that indicates the magnitude of mismatches in participants' movements. A substantial amount of research will be required to investigate the trade-offs that will have to be made between the amount of information provided, the sensory clutter that is produced, and the workload imposed on users in interpreting that information.

Finally, the advantage of the synchronized rule for symmetrical components of the task was not predicted. It highlights the advantages that could potentially accrue if trained and knowledgeable users are allowed to vary the techniques used to integrate their actions while they utilize a VE.

ACKNOWLEDGEMENTS

This work was supported by grant GR/L95496 from the UK Engineering and Physical Sciences Research Council. The work was carried out while R. A. Ruddle was employed in the School of Psychology at Cardiff University.

REFERENCES

- BASDOGAN, C., HO, C., SRINIVASAN, M.A., AND SLATER, M. 2000. An experimental study on the role of touch in shared virtual environments. *ACM Transactions on Computer-Human Interaction* 7, 443-460.
- BENFORD, S., GREENHALGH, C., CRAVEN, M., WALKER, G., REGAN, T., MORPHETT, J., AND WYVER, J. 2000. Inhabited television: Broadcasting interaction from within collaborative virtual environments. *ACM Transactions on Computer-Human Interaction* 7, 510-547.
- BOWMAN, D.A., AND HODGES, L.F. 1997. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proc. of I3D 97: ACM Symposium on Interactive 3D Graphics* pp. 35-38. New York: ACM.
- BOWMAN, D.A., JOHNSON, D.B., AND HODGES, L.F. 2001. Testbed evaluation of virtual environment interaction techniques. *Presence: Teleoperators and Virtual Environments* 10, 75-95.
- BROLL, W. 1995. Interacting in distributed collaborative virtual environments. In *Proc. of VRAIS 95: IEEE Virtual Reality Annual International Symposium*, IEEE, Los Alamitos, CA, 148-155.
- FRÉCON, E., AND STENIUS, M. 1998. Dive: A scalable network architecture for distributed virtual environments. *Distributed Systems Engineering* 5, 91-100.
- FRÖHLICH, B., TRAMBEREND, H., AGRAWALA, M., AND BARAFF, D. 2000. Physically-based manipulation on the responsive workbench. In *Proc. of VR 00: IEEE Conference on Virtual Reality*, IEEE, Los Alamitos, CA, 5-12.
- GREENHALGH, C., AND BENFORD, S. 1995. MASSIVE: A virtual reality system for teleconferencing. *ACM Transactions on Computer-Human Interaction* 2, 239-261.
- GOTTSCHALK, S., LIN, M., AND MANOCHA, D. 1996. OBB-Tree: A hierarchical structure for rapid interference detection. 5. In *Proc. of SIGGRAPH 96: ACM Conference on Graphics*, ACM, New York, 171-180.

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

- GREEN, M., 1995. The MR Toolkit Version 1.4 [Computer program]. Department of Computing Science, University of Alberta, Canada.
- GUIARD, Y. 1987. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of Motor Behavior* 19, 486-517.
- HAND, C. 1997. A survey of 3D interaction techniques. *Computer Graphics Forum* 16, 269-281.
- HINDMARSH, J., FRASER, M., HEATH, C., BENFORD, S., AND GREENHALGH, C. 2000. Object-focused interaction in collaborative virtual environments. *ACM Transactions on Computer-Human Interaction* 7, 477-509.
- HOLLERBACH, J.M. 2002. Locomotion interfaces. In *Handbook of Virtual Environments*, K. M. STANNEY, Ed. Lawrence Erlbaum, Mahwah, NJ, 239-254.
- HUBBOLD, R., AND KEATES, M. 2000. Real-time simulation of a stretcher evacuation in a large-scale virtual environment. *Computer Graphics Forum* 19, 123-134.
- LENGYEL, J., REICHERT, M., DONALD, B.R., AND GREENBERG, D.P. 1990. Real-time robot motion planning using rasterizing computer graphics hardware. *Computer Graphics* 24, 327-335.
- LOFTIN, R.B. 1997. Hands Across the Atlantic. *IEEE Computer Graphics and Applications* 17, 78-79.
- Macedonia, M.R., Zyda, M.J., Pratt, D.R., Barham, P.T., and Zeswitz, S. 1994. NPSNET: A network software architecture for large-scale virtual environments. *Presence: Teleoperators and Virtual Environments* 3, 265-287.
- MARGERIE D.M., ARNALDI B., AND PLOUZEAU N. 1999. A general framework for cooperative manipulation in virtual environments. In *Proc. of EGVE 99: Eurographics Workshop on Virtual Environments*, Springer, New York, 169-178.
- MASLIAH, M.R., AND MILGRAM, P. 2000. Measuring the allocation of control in a 6 degree-of-freedom docking experiment. In *Proc. of CHI 00: ACM Conference on Human Factors in Computing Systems*, ACM, New York, 25-32.
- MCNEELY, W.A., PUTERBAUGH, K.D., AND TROY, J.J., 1999. Six degree-of-freedom haptic rendering using voxel sampling. In *Proc. of SIGGRAPH 99: ACM Conference on Graphics*, ACM, New York, 401-408.
- RUDDLE, R.A., AND JONES, D.M. 2001a. Manual and virtual rotation of a three-dimensional object. *Journal of Experimental Psychology: Applied* 7, 286-296.
- RUDDLE, R.A., AND JONES, D.M. 2001b. Movement in cluttered virtual environments. *Presence: Teleoperators and Virtual Environments* 10, 511-524.
- RUDDLE, R.A., SAVAGE, J.C., AND JONES, D.M. IN PRESS. Evaluating rules of interaction for object manipulation in cluttered virtual environments. *Presence: Teleoperators and Virtual Environments*.
- SALLNÄS, E., RASSMUS-GRÖHN, K., AND SJÖSTRÖM, C. 2000. Supporting presence in collaborative environments by haptic force feedback. *ACM Transactions on Computer-Human Interaction* 7, 461-476.
- SHEPARD, R. N., AND METZLER, J. 1971. Mental rotation of three-dimensional objects. *Science* 171, 701-703.
- SONNENWALD, D.H., BERGQUIST, R.E., MAGLAUGHLIN, K.L., KUPSTAS-SOO, E., AND WHITTON, M.C. 2001. Designing to support collaborative scientific research across distances: The nanoManipulator environment. In *Collaborative Virtual Environments*, E.F. CHURCHILL, D.N. SNOWDON, AND A.J. MUNRO Eds. Springer-Verlag, London, 202-224.
- SLATER, M., SADAGIC, A., USOH, M., AND SCHROEDER, R. 2000. Small-group behavior in a virtual and real environment: A comparative study. *Presence: Teleoperators and Virtual Environments* 9, 37-51.
- SLATER, M., STEED, A., CROWCROFT, J., WHITTON, M.C., BROOKS JR. F.P., AND SRINIVASAN, M.A. 2001. Final report of the Collaboration in Tele-immersive Environments project. Retrieved from the World-Wide Web March 19, 2002. <http://www.cs.ucl.ac.uk/research/vr/Projects/Internet2/Report/finalreport.pdf>
- WIDESTRÖM, J., AXELSSON, A., SCHROEDER, R., NILSSON, A., HELDAL, H., AND ABELIN, A. The collaborative cube puzzle: A comparison of virtual and real environments. In *Proc. of CVE 00: ACM Conference on Collaborative Virtual Environments*, ACM, New York, 165-171.
- WANG, Y., AND MACKENZIE, C.L. 1999. Effects of orientation disparity between haptic and graphic displays of objects in virtual environments. In *Proc. of INTERACT 99: IFIP Conference on Human-Computer Interaction*, IOS Press, Amsterdam, 391-398.
- WARE, C., AND ROSE, R. 1998. Real handles, virtual images. In *Proc. of CHI 98: ACM Conference on Human Factors in Computing Systems*, ACM New York, 235-236.
- WICKENS, C.D., AND PREVETT, T.T. 1995. Exploring the dimensions of egocentricity in aircraft navigation displays. *Journal of Experimental Psychology: Applied* 1, 110-135.
- ZHAI, S., AND MILGRAM, P. 1998. Quantifying coordination in multiple DOF movement and its application to evaluating 6 DOF input devices. In *Proc. of CHI 98: ACM Conference on Human Factors in Computing Systems*, ACM, New York, 320-327.

APPENDIX A

Two MPEG videos, illustrating trials in the offset and C-shaped VEs, can be accessed from the Web page <http://www.comp.leeds.ac.uk/royr/video.html>. In each video, the view seen by Participant 1 (controlling the virtual human that is wearing the blue hat) is shown on top, and the view seen by Participant 2 (virtual human wearing the red hat) is shown

ACM Transactions on Computer-Human Interaction, 9, 285-308 (2002)

below. The white line seen on the floor at the end of each video is the finish line. Due to an oversight in data recording, the wireline feedback that participants saw while manipulating the virtual object is not shown in the videos. Neither video contains sound.