6.1 Introduction

The previous chapter described various techniques to support object-focused interaction for desktop CVEs. In concert, these mechanisms created an interface for the Furniture World task. The interface was designed to reduce the problems of working with objects that were observed in the exploratory evaluation: through supporting awareness of the shared environment; awareness of others’ actions within that environment; and navigating with respect to those actions. Interfaces and representations were described and designed such that object-focused work could more easily be achieved.

These designs were informed by the exploratory evaluation outlined in chapter 4. However, it still remains to assess the effect that these designs have on users’ ability to undertake collaborative work with virtual objects – and in particular the Furniture World task. Furthermore, these issues must be assessed for the possibility of extrapolating them into more general improvements to desktop CVE interfaces. This chapter undertakes such an analysis by presenting a targeted evaluation as described in chapter 3. To re-cap briefly, the targeted evaluation will highlight each interface technique in turn, to consider some aspects of their use in Furniture World. Following this, the ways that the overall system is used will be examined in more detail.

In the first section, the collection of data and the form of the targeted experiments will be compared and contrasted with the exploratory evaluation. Secondly, general observations regarding the task will be discussed. Thirdly, the interface designs as described in the previous chapter will each be targeted in turn. Some of the ways in which they manifest themselves as resources for object-focused interaction between participants will be shown. Fourthly, further data will be presented to highlight two of the key unanticipated results of using these designs. Finally, the evaluation will be
summarised, and its consequences for CVE interface design will be drawn out, in lieu of further discussion in the next chapter.

6.2 Conducting Furniture World II

In order to undertake the targeted evaluation of the new interface, it was necessary to collect data of participants performing the same task as for the exploratory trials. Apart from the use of a different interface, many of the aspects of the original experiments were retained.

- The task descriptions presented to participants were exactly the same.

- Data was collected in an identical manner – videos were obtained from the physical environment and the screen of each user, in addition to a four-picture shot of all views. Audio from the ‘real world’ cameras contained only sound from that participant’s physical space. Audio on the desktop screen video contained that participant’s voice in real time, and the other’s voice (as heard after a small transmission delay over the network). Participants were also interviewed, with the interview schedule as before. (Again, interview data was not explicitly analysed in itself, but rather used by the author to get a sense of the trial).

- The trials still lasted around one hour, with approximately ten minutes allowed to get used to the system, half an hour to perform the task, and ten to fifteen minutes for the interview.

However, there were also some differences between the first and second sets of trials (other than the redesigned interface).

- There were no trials with three participants, as the exploratory analysis had, in the main, highlighted issues which were relevant for two-party interaction, and had not concentrated on the differences arising when a third user was introduced.

- The participants themselves were different (except for one pair who had also performed in an exploratory trial). However, the users were broadly similar to
those from the first analysis: in occupation (mostly students), in gender (slightly more male than female), in previous acquaintance (a range from friends to strangers), and in number (six trials of two participants were performed).

The description of these similarities and differences is not to imply that the experiments were ‘controlled for variables’, in the sense of a traditional psychological experiment. Indeed, the trials were deliberately not controlled – in order to let interesting issues in the use of the new interface arise in participants’ interaction. However, these similarities are provided in order that the reader can derive a context for the following discussion as and when this chapter finds it useful to compare and contrast with data from the exploratory trials.

It may be worth reiterating that the experiments discussed in this chapter, indeed in this thesis as a whole, are not intended to show improvements in task performance, and are not assessed upon that basis. Rather, they are intended to gauge the intricate ways in which participants, given the interface resources provided, conduct object-focused interaction. This is for the purposes of extrapolating designs for other CVE tasks involving object-focused work; and also because this thesis has focused on the local achievement of organised interaction (rather than improving performance of the task). In the main, the targeted evaluation will consider participants’ orientations to the redesigned interface as it stands, and the ways in which they display these issues to each other. Thus, direct comparison can be seen only as a vehicle for describing how users orient to the different interface techniques. Contrasting of the two sets of trials, where necessary, will be on the basis of collaborative practices, and the ways that participants employ the interface resources. To begin with, however, some general observations regarding the participants’ accomplishment of the task will set the scene.

### 6.3 General observations

Participants accomplished the second set of trials fairly simply and, as with Furniture World I, were able to finish the task within an hour. Informally viewing the video
data gives an impression of the ease with which participants accomplish the fairly difficult task – which is not to say that users do not encounter troubles along the way, but rather that the interface seems to allow those users to embark upon apparently ambitious collaborative tasks in a relatively mundane fashion. It appears that participants in these trials are often able to produce virtual actions and rely upon others’ ability to see them; and that this occurs enough that participants’ talk consistently centres on the task itself, rather than the ways in which it is being accomplished.

However, these kinds of comments are sterile in the face of a detailed evaluation of the organisation of action that occurs within the data. This level of intricacy allows that taken-for-granted aspects of the interaction in Furniture World II may be revealed in order to uncover the relationship between the new CVE system resources and the participants’ action and interaction.

Thus, the following five sections analyse the collaborative work employed using the redesigned interface. The evaluation is separated into two parts. The first part (sections 6.4, 6.5 and 6.6), will target the three interface techniques employed in turn – peripheral lenses, extended actions, and target-based navigation. These techniques will be discussed in a preliminary fashion, before the second part (sections 6.7 and 6.8) looks in more detail at two key issues arising in the use of the Furniture World II interface. To begin the targeted evaluation, the first interface technique to be selected is the use of peripheral lenses.

6.4 Supporting awareness with peripheral lenses

Peripheral lenses were introduced into the Furniture World interface to increase the user’s horizontal field-of-view. The increased view was implemented to allow increased ‘peripheral awareness’: both of the surrounding environment and of the visual aspects of other’s actions with respect to that environment. These two issues will be dealt with in turn in this section, beginning with the ability of the lenses to reveal a peripheral environment.
6.4.1 Increasing awareness of environment with peripheral lenses

Whilst the use of peripheral lenses within the Furniture World environment gives participants an idiosyncratic view on the virtual environment (due to the distortion effects), it was anticipated that this extension would give a better sense of the world to participants, and of the objects it contains. Example 6.1 provides an illustration of the ways participants use this increased horizontal view of the world. Harry is engaged in searching outside the room for objects which have been accidentally discarded through the walls. He has left Fred inside the room to continue arranging furniture – thus they are in auditory but not visual contact.

Example 6.1 (Perspective: Fred)\(^{25}\)

Fred: found anything yet?

Harry: (0.2) errm, might have done (0.3) (can’t tell what that) is (0.4) I th- I think I’ve found the Hi-Fi

Fred: oh

Fred poses his question just as Harry turns to his left. At this point, the Hi-Fi appears in Harry’s left peripheral lens as a black blob\(^{26}\) against the background of the outside wall of the room. The object’s appearance coincides with the beginning of Harry’s utterance “errm, might have done”. He then continues to move his view around until the front of the Hi-Fi is centred in his peripheral view, and then place his mouse cursor over the object (compare the position of the mouse cursor, central in figure 6.1

---

\(^{25}\) As mentioned in earlier chapters, the notation for transcripts of talk is described in Appendix B.

\(^{26}\) The Hi-Fi can be difficult to identify, as it is entirely black from behind and has features only on the front.
and over the stereo in figure 6.2). At this point, he utters “I think I’ve found the Hi-Fi”.

The peripheral lenses are often used in this way, a key resource in enabling objects to be seen in peripheral areas. Harry’s search around the outside of the room is enabled through the ability to navigate parallel to the walls whilst keeping them in his peripheral view.

This is typical of interaction occurring inside the room as well. For example, participants may often discuss features of the environment, such as items of furniture,
without the need to rotate and find them – instead identifying them through the use of
the peripheral view. In Example 6.2, Harry and Fred discuss the bookcase, and its
position in the room.

Example 6.2 (Perspective: Fred)

Fred: oh: wh::at’s that book (. ) case doing by the
doors (0.8) is it by the door?

Harry: err (0.8) I don’t believe so

Fred: no [its ]-

Harry: [its not (0.4) its not by the door its
perspective

Fred: OK

Figure 6.3 – Fred: “What’s that bookcase doing by the door?”
The difficulty in seeing the spatial relationship between door and bookcase seems to arise because depth perception is not particularly good in the 3D view, which lacks stereoscopic functionality. The bookcase and door are along the same line of sight for Fred (figure 6.3), and his depth cues only come from the relative sizes of the two. However, Harry has the bookcase and door available in his left peripheral lens from an approximately orthogonal viewpoint (figure 6.4). Despite the distorting effects of the lens, this allows him to comment upon the spatial relationship between the two, and whether the objects are, in fact, close or far apart. The increased views offered by the peripheral lenses give participants a peripheral sense of the environment, and often allow them to see objects without the necessity of engaging with the problems of moving their view.

6.4.2 Increasing awareness of others’ actions with peripheral lenses

Peripheral lenses were intended not only to reduce problems with awareness of the local environment, but also to render visible others’ actions with that environment. Thus, it was hoped that the object-focused actions of others would be more easily seen and identified.
Informally it is relatively simple to see that the inclusion of peripheral lenses helps with awareness of actions as well as environment, but it is very difficult to find instances of interaction in which these features are explicitly orientated to. Certainly, the use of peripheral views may contribute to a visual awareness of actions within Furniture World II. Users will often arrange their avatars so that the other is visible in their peripheral lens, presumably to get some sense of each other’s location and potential actions. However, it is difficult to separate this peripheral awareness from the other mechanisms provided at the interface. As discussed in the next section, there are few instances where the visual extension of action did not become the key aspect through which peripheral actions were seen.

Moreover, the peripheral glancing facility, which was initially provided to allow peripheral actions to be viewed as undistorted, was utilised very little in Furniture World II, indeed only in a few instances over the course of all the trials. This was particularly surprising, as other data that the author has collected in CVE tasks shows a consistent use of the glancing facility. For example, in experiments described elsewhere, the peripheral glance was highlighted as a key resource in viewing others’ actions. Specifically, the ‘Duplication World’ experiments published in (Fraser et al., 1999) show how peripheral glancing is used to identify the actions of others. For example, the following fragment taken from that study shows the use of the peripheral lens to identify the location of a pointing gesture. This is followed by the use of peripheral glancing to reveal the details of the gesture at a distance.

*Fragment 1 (Quoted in Fraser et al., 1999)*

Gerry: I’ll go down to the far end of the red one umm and tell you sorta whereabouts they need moving

Bob: yeh umm d’ya see this pile I’m pointing at there?

Gerry: (2.0) ((G glances left)) oh yeh

Bob: they need to be moved sorta to my left so I’m just gonna
Gerry uses the peripheral lens to reveal Bob’s location. He subsequently glances to remove distortion from the lens, such that he can identify the object being pointed at ("this pile"). The peripheral lens is used to view peripheral action, and the glancing facility is subsequently used to resolve mutual orientation to the specific object in question.

There are a number of differences between these two experiments, which disallow any kind of detailed comparison to take place. However, it may be postulated that a key interface difference between the two trials is that Duplication World did not provide any kind of navigation based upon the actions of others. Informal observation of the two sets of data implies that, where peripheral glancing might have been employed, the ‘Other button’ is often pressed. Speculatively, this may be because actions cannot be performed with the mouse in a peripheral lens, even when undistorted through glancing. If the Other button focuses the participant’s main view on the action, they immediately have the ability to engage with actions of their own around the same area. These instances will become clearer in the forthcoming discussion of the use of the Other button. Thus, peripheral lenses may be key to a general peripheral awareness. However there are some problems in engaging with that periphery in order to act upon or display orientation to the object at hand.

6.5 Supporting awareness with extended actions

Three main actions provided by the interface that have been extended into the environment are grasping, pointing and looking. Grasping an object is shown by two arms extending out to touch the object and its becoming wire-framed during the movement. Users can point by extending a single arm to touch the object in question (elongating is either performed ‘manually’ or by the system). ‘Looking’ is portrayed in the environment through outlining the view with a semi-transparent frustum. In
some sense, all three actions have been modified so as to more explicitly display information about one’s actions to other users. In the next three sections, some examples will be presented to show the ways in which participants use these resources to accomplish their work with virtual objects. These will be looked at in order of grasping, pointing and looking.

6.5.1 Supporting awareness with extended grasping

The visible representation of grasping is to make the object a wire-frame for the duration of its movement, combined with the avatar’s arms both stretching to touch that object. The small fragment of talk in Example 6.3 provides an example of the use of these representations. This example is preceded by a long silence whilst Harry and Fred are moving chairs around. Harry then asks a question.

Example 6.3 (Perspective: Fred)

Fred: where you puttin that table
Harry: mm (0.7) over here

Just prior to his question, Harry’s view shows Fred moving the table. As Harry himself picks up a chair, the table turns into a wire-frame and begins to move. Fred’s arms have extended to touch it. Fred’s avatar is barely visible in Harry’s left peripheral lens (figure 6.5) and his arms appear in the left of Harry’s main view (figure 6.6). As the table is (nearly) central in Harry’s main view, these combined representations cause a noticeable change to the world (compare figures 6.5 and 6.6). The data suggest that it is these representations that unproblematically occasion his question. Indeed, Fred’s action is seen as, and see-able as, ‘moving the table’ such that Harry asks Fred to clarify the particulars of his actions with respect to the task – where he intends to put it.
The extended representations that denote grasping and moving an object are used as resources for collaboratively organising their placement. In the previous Furniture World trials, the thin connecting line could be visually indistinct, and confuse users regarding the actions of others.

However, not only are these problems often overcome in the new data, these representations can actually visually clarify conversational misunderstandings. This phenomenon is in contradiction to the norm of the earlier trials, in which conversation
was most often relied upon to clarify confusion regarding the visual or visible nature of on-going actions.

For example, consider the following example, 6.4. Kylie and Jess are deciding which corner to place an object in – specifically, which plug socket (“there’s one the other end of the room”) should be used.

**Example 6.4 (Perspective: Kylie)**

Kylie: Well I’ll put it in the corner then (0.3)

Jess: all [right

Kylie:    [cos there’s one the other end of the room

(. ) so (. )

Jess: yes

Kylie: (. ) [turn round ya bastard

Jess:    [well I’ll just (. ) yep (0.4) oh you’re

talking about the stereo I’m talking about the

television. Do we put them all in the same area?

Figure 6.7 – Jess’s initial view
Jess replies “yes” to Kylie’s suggestion with respect to the placement of the object. Kylie then begins to grasp the object, rotating it towards an orientation parallel to the proposed wall. (Her profanity is probably caused by difficulties in using the object rotation mechanism – the stereo itself seems to be the ‘bastard’ although, as an aside, this talk also concurrently displays these problems to Jess).

The representation of grasping indicates and highlights the stereo in Jess’s view (compare figures 6.7 and 6.8 where the arrows indicate). This wire-frame representation appearing in Jess’s view immediately precedes Jess’s comment that Kylie is talking about the stereo, not the TV. The representation of grasping highlights the referent of their talk, and causes Jess to clarify the object (“it”) previously discussed when the wire-frame appears (figure 6.8).

Thus, the visual representation of grasping the object enables Jess to pursue a course of repair regarding which object is being talked about. These occurrences stand in stark contrast to a reliance on talk as a repair mechanism for problems with visual action. Here, conversely, Jess and Kylie rely upon the representation of grasping to clarify their conversational misunderstanding.
6.5.2 Supporting awareness with extended pointing

The interface supports pointing in two ways. If users are pointing at an object with a definite position (e.g. a piece of Furniture), their arm will be extended to touch the object, thus making it more visible to others within the environment. If the user points to other areas that have no definite position (a corner, ‘the middle of the room’ etc.), then they are able to extend and retract the pointing arm manually.

There are many cases where pointing works well; and many cases where it doesn’t. For now, some of the simple ways in which extended pointing is used will be discussed. However, it should be noted that further examples, particularly those in section 6.8, will discuss some problems in using the extended actions (and pointing most prominently).

Informally, the extended pointing arm is relatively easy to see by others. Pointing is perhaps even more easily identified, because the extended arm may overcome some of the perspective problems with the 3D view in following the trajectory of a gesture. The closer an arm approaches the referent, the easier the trajectory of the arm, and hence the object, is to discern.

Additionally, there are common instances of participant’s avatars being out-of-view and yet objects being identified due to the appearance of the arm in-view. Thus, the extended pointing mechanism can be used as a resource in seeing referents when the other’s avatar cannot be easily seen. For example, the following excerpt involves Arnie and Steve trying to place a chair in which to sit and watch TV. We begin with an utterance during which Arnie extends his pointing arm towards a location on the floor.

Example 6.5 (Perspective: Arnie)

Arnie: stick stick the um (0.5) stick the chair (1.0)
    over (0.4) t- there

Steve: over?
Arnie: bit further rou- bit further in cos at the moment its miles away from the TV

Figure 6.9 – Steve’s view of Arnie’s arm

The first thing to note about this fragment is that the pointing arm is used successfully – to denote a candidate location as a feature of the environment, and to direct another to place the object at that location. Additionally, this is done whilst Arnie’s avatar is out of view for Steve, although he sees the arm in view, and where it is indicating (figure 6.9).

Another feature to note about this example is the pauses, ‘umm’s and vocal repetition which Arnie uses before and during pointing in the first line. These are typically produced while participants change from navigation mode to pointing mode, and in the case of them manually extending their arm, during this process. Other characteristics preceding and accompanying the production of pointing gestures include the use of phrases like “hang on, I’m going to point” (more examples will show this later in the chapter). These kinds of utterance are used to display to the other that ‘pointing’ as a process has begun, but that the gesture is ‘on its way’ or ‘not residing at its final destination’. These instances can be compared with
occurrences in Duplication World (Fraser et al., 1999), in which talk produced while extending a pointing gesture can be seen as intrinsically linked to the continued movements of the gesture, and for providing relevance of those movements for the other.

However, the difficulties with producing pointing gestures in Furniture World II run more deeply than this. This is not to say that pointing is not achieved and achievable for participants to show one another objects in the world; rather that the way pointing often manifests in interaction shows a key design assumption about the way that virtual environments are structured – in other words, about what the designer assumes in advance is a relevant ‘object’ for the task, and what is not. This issue will be discussed at length in the forthcoming section 6.8. For now, this section proceeds to investigate the use of the extended looking representation.

6.5.3 Supporting awareness with extended looking

The previous extended representations discussed have been extensions or ‘improvements’ on the first Furniture World system. Similarly, although less obviously, the explicit representation of the user’s field-of-view in the environment is an extension. It is intended to ‘improve’ the usual indicator of view for the represented user – showing orientation on the avatar (in the case of Furniture World I, and in many other CVEs, by giving the avatar a recognisable face and body to indicate orientation). The visible view frustum employed in Furniture World II extends this representation into the virtual environment, such that another’s view might be seen even though their avatar is not. The extent of the user’s (undistorted\textsuperscript{27}) field-of-view is explicitly outlined. Example 6.6 shows how, given the new frustum representation, the other’s general view is still available when the avatar is in view. Steve is positioned with his ‘back’ to the fireplace. Arnie raises the topic of the fireplace in a prolonged discussion about the layout of the room.

\textsuperscript{27} As mentioned and reasoned in the previous chapter, the peripheral lenses’ extents are not outlined in the Furniture World II task.
Example 6.6 (Perspective: Arnie)

Arnie: is there a fireplace somewhere? (0.4) is it behin[d you?

Steve: [what? all right (.) sorry this mic isn’t talking) err theres a- aren’t you looking at it?

Arnie: yeh I think I am

Steve: yeh

Figure 6.10 – Arnie’s view of the fireplace and Steve

After some initial problems with Arnie’s microphone, Steve asks whether Arnie is, in fact, currently looking at the fireplace. At this point, Steve and Arnie’s avatars are ‘facing’ each other. Arnie’s view (figure 6.10) shows Steve’s avatar oriented towards him, and his view frustum disappearing to either side, and Steve’s view shows a similar picture. Arnie proceeds to confirm that he thinks he is facing the fireplace.
The other’s general view can often be discerned in a similar way from the avatar’s
genral bearing – although this is not unexpected, given that ‘human-like’
representations display bodily orientation. Interestingly, however, the depiction
extends the other’s view into the environment so that it is possible to discern
another’s position without recourse to fast and intricate gaze movement, and indeed
without recourse to the other’s avatar. Consider Example 6.7, a fragment of talk
which follows a long period of silence in which Fred and Harry have been moving
objects in separate parts of the virtual room. Fred’s avatar is now positioned behind
Harry, and they are both facing in approximately the same direction (i.e. Fred can see
the back of Harry’s avatar in his main view).

Example 6.7 (Perspective: Harry)

Fred: OK

Harry: you’re behind me

Fred: yeh, I’m gonna grab the white chairs and put ‘em
round the table

Harry: oh go (on then) (.) right button

Figure 6.11 – Harry: “you’re behind me”
At the initiation of the conversation, Fred’s avatar has not been in Harry’s view (either main or peripheral) for some time. However, Fred’s looking representation has intermittently appeared in Harry’s view as it extends across the virtual environment. At the start of the example, during Fred’s “OK” utterance, the left edge of Fred’s frustum is Harry’s main view. Subsequently, as Fred rotates, the right edge of his frustum appears in Harry’s right peripheral lens (see figure 6.11). At this point Harry says, “you’re behind me”. Harry uses Fred’s frustum as a successful resource in locating Fred’s position and (orientation), without having to re-orient his view to find Fred’s avatar. This then enables Fred and Harry to proceed with their course of action regarding the task at hand – grabbing and placing the white chairs around the table. This example shows how the view frustum is used as a resource in locating the position and orientation of others, even when their avatar is completely out of view. This enables the participants to not have to establish direct visual contact, but rather to use the extended resources for an on-going awareness of the other’s movements within the environment.

The previous example illustrates how the position of the other can be seen without the requirement of seeing their avatar. The extension of the looking representation into the environment means that others can discern a user’s view of the environment even when their avatar is not visible. However, the two previous examples have been fairly coarse in the ways that participants work out what the other can see. In addition to this, the representation also gives the opportunity to intricately co-ordinate one’s own actions with their visibility by others. In Example 6.8, Dave and Tom are discussing where to put the table and chairs. Tom has extended his pointing arm towards a wall by a corner of the room (see figure 6.12, Dave’s view prior to the beginning of the example). Dave has rotated his view from looking at Tom’s avatar by following the pointing arm to its destination.
Example 6.8 (Perspective: Tom)

Dave: yeh i- i- i see where you mean (0.4) right in the corner yeh?

Tom: well somewhere over here cos they're not really (.) that much use for much else but some people could sit (an eat out)

Dave: yeh (1.0)

Tom: play a game or something (0.8) cos the chairs and the table go together

As Dave and Tom engage in talk to agree the placement of the furniture, Dave has continued to rotate his view from looking at the arm towards looking at the table and chairs. Figures 6.13 and 6.14 show Dave and Tom’s respective views at this point. As Tom utters “go together”, Dave’s frustum passes the end of Tom’s pointing arm. Tom releases the mouse button, and his pointing arm drops. Compare their views again in figures 6.15 and 6.16. Tom is able to coordinate his pointing gesture with its visibility.
for Dave. Indeed, just as Dave uses Tom’s gesture as a resource to discuss the placement of the table, Tom uses Dave’s frustum as a resource to produce the gesture. The ability to discern the visibility of one’s own actions by others allows an intricacy unseen in the previous Furniture World Trials.

Figure 6.13 – Dave’s view as he rotates past Tom’s gesture

Figure 6.14 – Tom’s view of his own pointing arm and Dave’s frustum
This aspect of designing actions for their visibility by others is further demonstrated by the following example. It shows how the extended view frustum allows actions to be produced by considering whether the other can see the object of those actions. It begins with Harry asking which of the large chairs can be disposed of. Fred is trying to place the TV in the corner.
Example 6.9 (Perspective: Harry)

Harry: which do you think is a- best of these two
   chairs the big three seater or the (0.3) big one
   comfy one (0.6)
Fred:  (errl- l- let me) wait there (trying to) put the
   TV there
Harry: cos I don’t think we want this umm (0.3) this
   big seat that I’m moving now (.) I’ll move it
   into yer view so you can see it (0.7) oh: (.)
   sort of (1.0)
Fred:  hhh
Harry: do you think we need this? (0.5)
Fred:  need what
Harry: this- this one I’m selecting (0.4) its in yer
   view hhh
Fred:  well just leave it out (to) the way for the time
   being (0.5)
Harry: well no do you think we need it at all (0.3)
Fred:  we could do (0.3) prrr
Harry: that doesn’t help me
Fred:  well just (0.4) fffhhh ((sighs))
Figure 6.17 – Harry’s view at the start of the example

Figure 6.18 – Harry moves the “big one comfy one” into Fred’s frustum

Figure 6.19 – Fred’s view of Harry’s arm (circled) and the wire-framed chair
This example shows how participants can use the frustum to anticipate the potential for the other not to see an object. Harry can see the chair is not in Fred’s view (i.e. not within the bounds of his frustum, see figure 6.17). Therefore, in order to show it to Fred, he moves the chair into the frustum (and thus Fred’s view), so that they can discuss whether it can be disposed of (see figure 6.18). As well as the use of the frustum to show the object to the other, this fragment of talk nicely illustrates how the frustum renders its ‘owner’ accountable to attend to the object. The possibility of technical reasons for not seeing the object is reduced, because the frustum implies that what is contained within it can be seen. Visibility is no longer a potential get out clause for ignoring the object. It should be noted from figure 6.19, however, that the chair and Harry’s (circled) arm are not particularly obvious, despite its location in Fred’s main view. This is an issue that will be returned to in section 6.7.

The previous examples have shown that the course of object-focused actions can be transformed by using the frustum as a resource to anticipate the limitations of viewing. In effect, the view frustum is used for coordinating the visibility of action, such that it both supports and transforms the ways in which gestures, grasping and so on are accomplished within participants’ visual restrictions.

### 6.6 Navigating with respect to the actions of others

Navigation of one’s viewpoint around Furniture World II can be achieved in two ways. As is commonly provided for desktop CVE interfaces, users can move their view in an independent ‘cartesian’ manner: forwards, backwards, turn left and turn right. However, Furniture World II participants can also move their viewpoint based upon the actions of their co-participant. The simple algorithm outlined in the previous chapter, and presented at the interface through the ‘Other button’, allows the user to move his view to a target based upon the system’s determination of ‘what the other is doing’. This is intended to overcome some of the slow and limited ways of seeing what the other participant is doing at any particular moment.
Example 6.10 provides a good illustration of the ways in which target-based navigation is achieved in the trials. In this situation, Dave points at an object for Tom. Subsequently, Tom is able to turn to look at the object using the Other button.

Example 6.10 (Perspective: Dave)

Dave: can you see that now Tom?
Tom: I can’t (.) see it (hang on a sec) I can see you sort of looking at me (.) so I presu-
Dave: f (.) if you look at what I’m p- If you look at what I’m doing (2.6) should spin you
Tom: oh yeh I can see it now

Figure 6.20 – Dave’s view as he points to “it” (the bookshelf)
This fragment of talk occurs at a point in the trial where Tom’s machine has (for task-unrelated reasons) slowed down considerably, exasperating the clumsy ‘cartesian’ navigation method. This method is initially tried by Tom, who begins to rotate from the wall he is facing (figure 6.20), as he utters “hang on a sec”. The object Dave is pointing at is located to Tom’s right. Dave mentions the Other button as a way of getting Tom to look at the object: the object in question is ‘what he is doing’. The way that the CVE system defines the target of Dave’s pointing gesture means both that Dave’s arm extends, and the object is identified by the system as a navigation target for Tom to use. In other words, there is little difference (or at least enough similarity for their practical purposes) between the relevant object for the participants, and the target object that the system uses and provides for Tom’s interface. Tom presses the Other button, which rotates his view to centre on the set of shelves (note the movement of Tom’s avatar and view frustum between figures 6.20 and 6.21). In this context, system object and interactional object are adequately alike for the task at hand. As a result, Dave and Tom are able to repair difficulties in orienting to the object in question through producing and following a referential gesture.

There are cases, therefore, where interactionally relevant objects can be observed using target-based navigation. Tom and Dave use this form of navigation as a strategy to repair some problems in establishing mutual orientation. The next example comes

Figure 6.21 – Dave’s view as Tom turns using the Other button
from further into their trial, and shows how the Other button can be used as an aid in moving quickly between orientations, and perhaps more importantly, that this can be subsumed within the overall activity at hand. As we join the action, Tom has grasped the standard lamp, but is having some problems in placing it exactly, due to mouse over-sensitivity.

**Example 6.11 (Perspective: Dave)**

Dave: where’s that light goin?

Tom: (0.6) oop (.) not over there (1.0) (       )

Dave: def– I’d say we definitely need (.) a desk over by the socket (1.8) ((presses Other button))

Tom: ((tuts))

Dave: just watchin where ya ((Tom drops lamp)) putting the lamp I thought you said that corner

![Figure 6.22 – Dave’s view before pressing the Other button](image-url)
The resources that are employed by both participants to organise this problem are extremely intricate and yet tacitly interwoven with the topic at hand. So, as Tom is involved in the task of moving the lamp, Dave rotates to his left to view another part of the room, and comments upon a possible arrangement for that. He has the “desk (.) over by the socket” in his view, whereas Tom’s avatar and the lamp are barely visible in his right peripheral lens (figure 6.22). However, in the 1.8-second pause, Tom moves the lamp past Dave’s view, such that it appears and then quickly disappears. Dave then presses the ‘Other’ button, which automatically rotates his view to centre upon the lamp that Tom is grasping (figure 6.23). Dave then questions the placement of the lamp as compared with its proposed position that was discussed prior to this example.

The use of the ‘Other’ button enables Dave to quickly move between topics, both visually, but perhaps more importantly, with respect to the task. The fact that these different aspects are moved between, and that different places in the virtual environment are made relevant in the participants’ interaction, is subsumed within the subject matter under discussion – the placement of objects relevant to the task, in this
case the desk and lamp. In other words, target-based navigation is used to allow the
design task to remain the focus of the interaction, rather than that interaction itself
becoming a problematic topic.

The previous three sections have looked at each of the three interface techniques
discussed in chapter 5 – peripheral lenses, extended actions, and target-based
navigation – and mentioned some of the ways in which they are used in working with
objects in Furniture World II. The following two sections now proceed to look at two
unforeseen features of participants’ interaction in Furniture World II. The first,
discussed in the following section is the ‘subjectivity’ of the environment between
participants, relevant to two of these techniques. The second, looked at in section 6.8,
is the problem of object definition that the Furniture World system employs, and its
consequences for collaborative work.

### 6.7 Problems with awareness and subjective views

Subjectivity is a difficult issue in collaborative virtual environments. It has been
investigated for particular applications and in various conceptual ways (e.g. Snowdon
et al., 1995, Smith and O’Brien, 1998). A key issue arising in these investigations is
that information (e.g. 3D visualisations) can be presented in different ways to
different participants populating the same environment. Commonly ‘subjectivity’ is
treated as a property of the virtual environment, which might pose problems (or
provide useful features) for users. However, it is also clear that collaborative use of
virtual environments is inherently subjective. For example, elsewhere (Vaghi et al.,
1999, Fraser et al., 2000, Jordan and Ruhleder, 1999), it has been argued that
distributed communications technologies will always suffer from some aspects of
‘subjectivity’ due to the effect of network delays. Timing of updates to different users
views of the environment is subject to network lag that is, at the very least, limited by
the speed of light. In practical circumstances, delays will often be quite large due to
poor bandwidth and high traffic on current wide-area networks. These delays cause

---

28 Subjective in the technical sense mentioned, rather than the sociological sense, where its meaning is
a hornets nest of debate concerning ‘objective’ versus ‘subjective’ social participation and structure.
differences in the timing of changes to the world, and thus affect the causality of actions.

In the case of Furniture World II, however, another form of subjective views arises. This relates to the interface techniques outlined, which are designed to provide increased awareness of the environment, and others actions in the environment. In the next section, the first of these problems is discussed. This involves the spatial differences seen when users have different parts of the same environment distorted by peripheral lenses. The second problem is mentioned in section 6.7.2, and relates to the inability to see the effects that one’s own view frustum might have on the co-participant’s view of the virtual environment.

6.7.1 Peripheral lenses and subjective views

The first issue relating to a subjective Furniture World is related to the peripheral lens technique. Peripheral lenses provide distorted 60-degree views to either side of the user’s main view. Thus, it stands to reason that unless participants are in exactly the same location, and at exactly the same orientation, they will see different parts of the environment as distorted. Thus, problems may arise when two participants are viewing the same parts of the environment, and yet different regions are distorted for each.

It is common for such differences to be momentary in the Furniture World II data (perhaps because participants commonly centre their main view upon features of the environment under discussion). However, to illustrate this phenomenon, I shall refer back to Example 6.8 (quoted again below for convenience).

---

Example 6.8 (requoted) (Perspective: Tom)

Dave: yeh i- i- i see where you mean (.) right in the corner yeh?
Tom: well somewhere over here cos they're not really (. ) that much use for much else but some people could sit (an eat out)

Dave: yeh

Tom: play a game or something (. ) cos the chairs and the table go together

Figure 6.24 – Dave: “right in the corner yeh?”

Figure 6.25 – Tom: “well somewhere over here”
From Dave’s perspective, Tom extends his pointing arm to point “right in the corner” (see the distorted end of the arm in figure 6.24). However, Tom is standing next to a wall and, from his perspective actually points away from the corner to a position on the far wall. Dave asks for confirmation that Tom’s gesture points right in the corner. Indeed, his view of the gesture seems to show the gesture pointing to the corner, at least as a practically adequate approximation. This is due to the corner and the end of the gesture being contained within his left peripheral lens. Tom’s view suggests a different referent – he sees the distance between the gesture and the corner as much further because it is in a potentially undistorted condition. Thus he replies to Dave’s assertion with “well somewhere over here” (figure 6.25), simultaneously moving his arm around to further indicate the area in question.

Although the differences, in the majority of cases in the Furniture World II data do not suggest any great interactional difficulties for participants, it would not be too difficult to imagine situations or application where participants’ differing spatial separation of features of the environment might cause more problematic sequences. In particular, this might prove the case where relevant features of the environment to the task were relatively small and close together, thus increasing the spatially distorting effect upon distinguishing between them.

### 6.7.2 Extended looking and subjective views

Section 6.5.3 showed how participants orient to the extended looking representation, a visible view frustum, as a resource in determining how to coordinate actions with their visibility by others. However, a participant is not able to see his own frustum which provides others with information about his view. Thus the effect of moving one’s own view on the co-participant’s view of the world cannot be seen.

This is at odds with the fact that participants are often wary of blocking each other’s line of sight – especially in Furniture World, where:
• there is no solidity and thus participants can do things like walking through each other’s avatars

• there is no haptic feedback that an avatar collision has occurred (e.g. one participant is standing directly in front of another by reversing in front of them)

• it can be difficult to glance backwards to check another’s line of sight when they are out of view

Examples 6.12 and 6.13 show how, as a matter of course, Steve and Arnie organise sight-blocking movements. In the first example, Arnie anticipates blocking Steve’s line of sight to the object he is moving (the “desk thing”), and notifies him of the possibility.

**Example 6.12 (Perspective: Steve)**

_Steve: I’m gonna move that desk thing (0.4)

[out the way

_Arnie: [I’m just walking across your (0.4) line of sight (I’ll) be out the way in a minute_

Similarly, in Example 6.13, Arnie and Steve have positioned their avatars in proximal positions and at similar orientations. As they move in small increments, their talk negotiates whether they are in the way of each other.

**Example 6.13 (Perspective: Arnie)**

_Arnie: am I in your way?

_Steve: no I was just looking (.) am I in your way now?
However, despite these negotiations of the need to not obscure the other’s line of sight, ‘sight blocking’ occurs more than one might expect within the data corpus, due to the use of the view frustum.

The frustum projects exactly at the edges of the user’s view, and thus cannot be seen by that user. Participants can only see the frusta of others, and thus it can be difficult to determine the effect of moving one’s own view on theirs. In Example 6.14, Arnie and Steve discuss the placement of chairs around a table.

Example 6.14 (Perspective: Arnie)

Arnie: what have you done with these chairs?

((laughing))

Steve: what chairs ((laughing)) (1.0) what's wrong with em . I can't see jack (1.0) w- I can but (0.6) the old perspective things quite hard
Arnie asks about the layout of chairs (which has been changed by Steve since he last viewed the area). Steve turns to look at the chairs in question, chuckling. Steve says “I can’t see jack”, just as Arnie’s view frustum comes between his viewpoint and the table and chairs (figure 6.26). The semi-transparent effect of Arnie’s view movement causes Steve to mention, “the … perspective thing’s quite hard”.

Example 6.15 illustrates how Fred and Harry orient to this problem. Fred rotates his main view until it centres on Harry’s avatar. He then rotates to include the desk in his view.

*Example 6.15 (Perspective: Fred)*

Fred: what you d- d- what you doin with that desk
(0.7)

Harry: puttin it near a pink chair again (0.4) will you stop blinding me with that (.). flashlight

Fred: what flashlight?

---

29 It is likely that “jack” is an abbreviated version of the colloquialism ‘jack shit’, meaning nothing.
Harry: your [vision
Fred: [oh

Figure 6.27 – Harry: “Putting it near a pink chair”

Figure 6.28 – Harry: “stop blinding me with that flashlight”

If this example is seen from Fred’s perspective, all that can be seen is that he rotates his view with the edge of his main view near Harry’s avatar, and his mouse moving around the desk he has asked a question about. His reply to Harry “what flashlight?” seems perfectly reasonable. On the other hand, when the example is viewed from Harry’s perspective, Fred’s frustum appears across his entire view just prior to his
exclamation (compare figures 6.27 and 6.28). Perhaps more importantly, it bisects Harry’s avatar and the object he is grasping. The effect of Fred’s view movement on Harry’s world is unnoticeable and unavailable from his perspective. However, it makes definite changes to Harry’s view, to the extent that he asks Fred to “stop blinding” him. This suggests that participants can be unaware that their movements are causing significant changes in another’s view.

Example 6.12 and 6.13 showed how obscuring with the avatar is anticipated and mentioned. However, the visible view frustum is entirely invisible to its owner. Thus, it causes a subjective aspect to participants navigating their view around the virtual environment, such that the effect of this on others’ views cannot be seen.

This subjective obscuring effect becomes notably problematic where the frustum happens to obscure a wire-framed object (i.e. an object that is being grasped by the other participant). In any case, the wire-frame makes the object less visible for the duration of the grasp. An example of wire-framing conspiring with the visible view frustum to make objects less noticeable is provided in Example 6.9, a portion of which has been re-quoted here to help the reader.

Example 6.9 (partially requoted) (Perspective: Harry)

Harry: do you think we need this? (.)

Fred: need what

Harry: this- this one I’m selecting (.) its in yer view hhh

Fred: well just leave it out (to) the way for the time being

Harry: well no do you think we need it at all (.)

Fred: we could do (.) prrr
Harry: that doesn’t help me
Fred: well just (.) fffhhh ((sighs))

It will be recalled that Harry has moved a chair into Fred’s view frustum so that he can see it. However, the effect his own view frustum is having on Fred’s ability to see the object is not visible to Harry. Harry’s view extends between the chair and Fred’s location, and thus Fred’s view of the chair is obscured to some extent by Harry’s frustum. As Harry is “selecting” the chair (by grasping and releasing it to make it flash between a wire-frame and ‘solid’), the chair arguably becomes less visible to Fred behind the frustum (refer to figure 6.19). Whether Fred is being vexatious in ‘not seeing the chair on purpose’ is not at issue here, but rather that it is clear from data that what one participant sees as highly visible is, in fact, barely visible to the other. This happens without seeing the view frustum, and thus the effect that it has on the co-participant’s view. Therefore, the view frustum can hinder participants in designing actions to be seen by others.

6.8 Defining objects in virtual interaction

Having looked at some problems in maintaining intersubjectivity when using peripheral lenses and view frusta, it is possible to proceed to the second problem that arises – perhaps the most critical in accomplishing the task in Furniture World II. This section delineates the problem, and some of the ways in which it manifests itself through participants’ interaction.

Any CVE application structures the virtual world in systematic ways. In other words, the application (environment, world) is programmed and created by categorising its features. Furniture World II is built on top of the MASSIVE-2 system which, as with all CVEs systems, constructs the environment into spatial and behavioural arrangements. The simple structure of the Furniture World environment is shown below.
This structure applies equally to Furniture World I and II, and simply denotes that some graphical objects loaded into the system are not treated as ‘objects’ for the task at hand. An ‘object’ as defined by Furniture World is something that can be picked up and moved in the design task. At any one time, these objects have a position and orientation which may be changed by participants. They are describable in simple and singular ways – ‘a sofa’, ‘a bookshelf’ and so on. Background features however, whilst graphically visible in the environment, are not assigned ‘object status’. They are fixed for the task and it may be less easy to define their exact position mathematically. Indeed, the separation between ‘window’, ‘fireplace’ and so on above is only provided to give the reader a sense of what a background feature might be. There is no actual ‘window’ object, for instance, in the Furniture World system.

However, when the new interface techniques are employed – techniques that explicitly use definitions of ‘objects’ for enhancing awareness of actions around them – this distinction becomes more obvious than in Furniture World I. For example, the interface can now ‘extend the arm to touch the object’s position’ or ‘orient the view to centre on the object’s position’. This section describes some of the practical ways in which system definitions of ‘what is an object’ can cause problems to manifest themselves in the interaction of Furniture World II participants.
6.8.1 Object intricacy as a relevant feature of interaction

To begin to explore this issue, one way in which Furniture World structures features of the environment at odds with commonsensical notions is that of *intricacy*. In other words, in co-present interaction, participants often invoke *parts of* objects as relevant. What exactly an object is, is tacitly negotiated in collaboration. For example, we might discuss a door, its handle, hinge, paint-job and so on without worrying about what is being talked about in terms of intricacy, position, spatial extent and so on. Participants constitute the relevance of the reference through their on-going interaction.

Within the Furniture World data, it is clear that participants may be unsure as to the intricacy at which objects are defined. One good example is the bookshelf. It is structured by the CVE as a single ‘object’ for the purposes of position, orientation etc. However, it also presents a commonsense image of ‘books’ resting on ‘shelves’. As a result, participants often discuss the ‘objects’ at a different level of intricacy to the system’s ‘definition’. To illustrate this, consider Examples 6.16 and 6.17. Both Steve and Jess are trying the environment out at an early stage in their respective trials. They both orient to a book from the shelves by putting the mouse cursor over a particular book and in Jess’ case pressing the grasp button. However, the CVE system treats this as an attempt to grasp the entire bookshelf.

*Example 6.16 (Perspective: Steve)*

Steve: um (.) wicked (.) how about with books? nah it’s the whole thing isn’t it we move

*Example 6.17 (Perspective: Jess)*

Jess: hey:: (0.4) oh its picked up the whole damned thing. What does that one do? Lifts it

Kylie: yeh
Consider the second of these two examples. Jess places her mouse cursor (circled) over a book (figure 6.29) and presses the mouse button. The entire bookshelf turns to a wireframe, indicating it has been grasped (figure 6.30). This occasions her to say “oh its picked up the whole damned thing” as she begins to move the shelf around. ‘Books’ as a commonsense object in the world are at odds with the resulting acquired object of the grasp – the entire set of shelves. However, the grasping representation indicates the system’s reaction to the grasp, and allows Jess to discern the system’s
behaviour: the intricacy at which the object is defined. It may be that these occurrences usually occur at the start of the trials because participants start to navigate the system in order to discern its definitions of intricacy regarding ‘what an object is’.

### 6.8.2 Objects as undefined features of referential gestures

Many occasions of object-focused interaction are much more problematic than this. In particular, instances whereby users are trying to point to objects that are not defined as such by the CVE system. Such occurrences pervade the Furniture World II data, and highlight the problems of the use of the CVE system to pre-identify environment structure.

One way in which these difficulties arise is if the system cannot determine what is being pointed to. The next two examples are typical of problems with pointing at objects. The reader will recall that the representation of pointing is a single extended arm. The extending is either done by the CVE system (if the mouse is over a system-defined object) or by the user (if the mouse anywhere else, i.e. over a background feature). In cases where the mouse is over a background feature, the arm remains near the avatar, the pointing gesture defaulting to a ‘human length’ arm in preparation for extending by the user. Unfortunately, various factors combine to make it very difficult to see the arm at all: the limited vertical field-of-view; the use of a ‘through the eyes’ viewpoint; and the short arm length. Thus, participants often can’t easily tell whether their attempt to point has worked or not. This is in marked contrast to the grasping examples, in which it was immediately obvious how the system had reacted differently from what was expected. Confusion arises as to why the point has not occurred. In Example 6.18, Arnie and Steve are creating an area to watch TV in.

**Example 6.18 (Perspective: Arnie)**

Arnie: we’ll get the chair in anyway (   )
Steve: yeh (0.8)

Arnie: I reckon we want the chair [( )

Steve: [oh::: what’s it got?

Arnie: (1.0) we want the chair (0.4) about (2.0) point

(3.0) we want the chair sortuv around here

Steve: around where?

Arnie: facing towards the TV

Figure 6.31 – Arnie: “We want the chair sortuv around here”

Arnie changes to ‘pointing mode’ and says “I reckon we want the chair…” Steve is trying to move another object, but apparently grasps the wrong one (“what’s it got?”). As described earlier in the chapter, Arnie’s next turn consists of a number of pauses, interspersed with talk which displays that his point is ‘in progress’. Here, he has pressed the mouse button to raise his arm over the course of this turn, but subsequently releases it. The avatar arm has not appeared, being too short to intrude into his undistorted field-of-view. Eventually, after a two-second pause, he says ‘point’ (as if instructing his virtual proxy), and presses the button again, moving his
mouse around. Finally the pointing arm slightly appears in his view, and he then begins to extend it towards the candidate location on the floor (figure 6.31). However, over the course of the troubles, Steve has begun to move the chair into the approximately correct area. Thus he displays his inattention to the eventually produced gesture. The accompanying gesture is not treated as sufficient to disambiguate what is “sortuv around here”, and Steve says “around where?”. Thus, we can see that the difficulties in the temporal production of pointing to ‘non-system objects’ are not only a problem for the user, but also for the participants for which the gesture is produced.

In Example 6.19, problems are exacerbated by the very fact that the other’s avatar is in view. In this case, Steve and Arnie are discussing candidate locations for the TV. Steve has already mentioned that the TV must go by a plug socket. As we join the action, he is beginning to go through the process of producing a pointing gesture towards one of these sockets. The first comment is a typical one in the use of the pointing facility – “right I’m gonna point”.

Example 6.19 (Perspective: Steve)

Steve: if you look (.), right I’m gonna point (.).

Arnie: if you point an then I’ll d- I’ll look at what you’re doing

Steve: which one’s point (.), am I pointing? (0.8)

Arnie: are you pointing?

Steve: its the (.), left-hand button isn’t it (.)

Arnie: um:: (0.4) yeh I think so

Steve: OK hang on yeh (it works) (.), need to ex-

[ extend (.)]
Arnie: [hang on (                   )
Steve: oh
Arnie: OK go on point away
Steve: right yeh (.) somewhere along there
Arnie: yeh

Figure 6.32 – Steve: “Am I pointing?”

Figure 6.33 – Arnie: “Are you pointing?”
This example contains many similarities to the previous one – Steve’s pointing arm repeatedly appears and disappears from Arnie’s view (figure 6.33). There is only a small part of the arm, barely perceptible in Steve’s peripheral view as he points (figure 6.32). Critically, however, not only does Steve encounter problems in producing a pointing gesture towards a ‘background feature’ (“am I pointing?”), but this difficulty causes problems for Arnie about whether the repeated gestures are in the course of being organised for his benefit, accidentally, problematically or whatever (“are you pointing?”).

The conversational differences between instances where system and talk objects coincide for all practical purposes and where they differ are marked. When the feature being pointed to is a background feature, the user must extend the pointing arm manually, and this may often take some time, or be unrecognisable to the pointer, who cannot use the Other button to turn to the object (more on this later). These differences mean that a pointing gesture has an unpredictable spatial and temporal organisation. This is a problem for both the person pointing and, critically, for the other participant.

In Example 6.20 (which extends Example 6.11), the limited extension of the pointing gesture done by a participant misleads the other about what is being pointed at. The spatial organisation of the gesture when the ‘extending’ is ‘incomplete’ will often not adopt the same precision as an instance where the system extends the arm to exactly touch an ‘object’. It will be recalled that Dave and Tom are discussing the placement of the lamp. The candidate location for the lamp is one corner of the room.

Example 6.20 (Perspective: Dave)

Dave: where’s the light goin?

Tom: (0.6) oop (.) not over there (1.0) (          )

Dave: def- I’d say we definitely need (.) a desk (.)

over by the socket (1.8) ((presses Other button))
Tom: ((tuts)) (1.0)

Dave: just watchin where ya ((Dave drops lamp)) putting the lamp I thought you said that corner

Tom: I’m trying to go for the (0.6) ((Tom changes to pointing mode, Dave picks up lamp)) hhh trying to go for the corner over there somewhere=

Dave: =have we both got hold of this light here

Tom: no I’m p- tryna point (at) the corner over there (0.8) that’s where I meant (.) or did you mean the other corner

Dave: that corner? (0.3) is that the one you meant

Tom: I meant the other corner where the sockets are

Figure 6.34 – Dave’s view as he puts the lamp in a corner
Tom has encountered troubles in grasping and moving the lamp to the ‘correct’ place. Dave grasps the lamp after Tom has dropped it to point at the “corner over there”. Tom extends his pointing arm approximately half-way to the corner (the corner not being a system object\textsuperscript{30}), saying “that’s where I meant” (figure 6.35). The trajectory of Tom’s arm looks vastly different in Dave’s view to Tom’s view, especially with difficulties in depth perception in the virtual space, such that Dave identifies the wrong corner. It is easy to see why when we compare figures 6.34 and 6.35 (the lamp is circled in both figures). Had the corner been a system object, Tom’s pointing arm would have instantly extended to touch it. Yet the extended arm is seen and see-able in Dave’s view as pointing to an entirely different corner. Just which feature of the environment Tom really means is sufficiently ambiguous in both talk and gesture that the wrong candidate location for the lamp is selected.

If the corner were to be defined as a system object, then the ambiguity in gestural reference might well be resolved. Yet the problem from a systems design perspective in this case is how the position and extent of a ‘corner’ is defined. For the CVE system to achieve this definition, it must be divorced from its relevance produced in the context of the participants’ interaction. Imagine the variety of different sizes,}

\textsuperscript{30} Tom’s arm actually extends to touch the sofa object, and then he moves it upwards at this length to point at the corner. This practice is given further consideration in the next section.
positions and locations that “in the corner” can mean in any given situation in Furniture World.

**6.8.3 Strategies for referring to system-unaware objects**

One strategy that participants often use for coping with the aforementioned difficulties in pointing is to use system-defined objects as ‘proxies’. Indeed, in the previous example, Tom uses the sofa, which is about half-way to the corner, as a proxy object. In these cases, there may be a system-aware object in a similar direction to the required background feature. The object may be used to instantaneously extend the arm in approximately the right direction. Talk is then used to constitute the adequacy of the gesture, by naming the referent. In Example 6.21 below, Steve points to a chair as a system-defined object that is on the trajectory of the corner he describes.

*Example 6.21 (Perspective: Arnie)*

Arnie: We’re obviously going to have to have a little desk ah as well cos we’ve got a little filing cabinet and a desk (0.8) we’ll have to find a little area for that

Steve: (0.8) I reck- you don’t think we have the dining table right more in the corner (0.6)

Arnie: errm

Steve: and then we can put the lamp (0.4) errm i:n (0.4) hang on I’m going to point (0.4) i::n (.) that corner over there (0.7) can you see where
I’m pointing

Arnie: yeh but theres no socket over there

Steve: (0.8) arse

Figure 6.36 – Arnie’s view of Steve’s pointing gesture

After his (typical) utterance, “hang on, I’m going to point”, Steve’s arm extends out to touch the chair that is between him and the corner. His arm remains pointing at the chair, and yet his talk invokes the referent as “that corner over there”. Arnie can see the extended arm, chair and corner, and proceeds with the task at hand, by pointing out a problem with having the lamp in that corner – that there is no socket there. Whilst there are, as seen in previous sections, difficulties in producing the pointing gesture in the first place, the instantaneous extension of the arm to touch a system object allows the talk to subsequently proceed back to the task at hand. Arnie constitutes the corner referred to in Steve’s talk via his gesture as practically adequate as a candidate location, such that he can proceed with the business at hand – refuting that location due to the lack of plug sockets there.

In the following Example, 6.22, Steve uses a chair proxy in a similar manner to the way he did in the previous example. However, in this case, he also moves the now-
extended arm around to the correct background feature to be indicated, and ends up performing an intricate sweeping referential gesture.

**Example 6.22 (Perspective: Steve)**

Steve: right (0.4) OK I’ve got a plan  
Arnie: right  
Steve: put ermm (0.5) we could put the TV (0.7) like  
(0.8) hhh can you see my hand? (0.6) Oh no  
[hang on (0.4)]  
Arnie: [ermm yeh=]  
Steve: =(sortuv) on that wall (.) or there  
Arnie: yeh  
Steve: an then we could have sofas coming (0.8) err (.)  
	sortuv out  
Arnie: yeh thats [what I was thinking  
Steve: [( ] out (0.4) like  
that  
Arnie: yeh (.)  
Steve: *hey* is that quite a good point thing?  
Arnie: yeh thats quite a good point (.)  
Steve: cheers mate ((laughs)) good sweep  
Arnie: ((laughs))
Figure 6.37 – Steve: “we could put the TV … on that wall”

Figure 6.38 – Arnie’s view as Steve says “we could have sofas coming …”

Figure 6.39 – Arnie’s view as Steve says “sortuv out … out like that”
This example consists of two pointing gestures by Steve which use system objects as proxies. Firstly he points to the prospective location of the TV, using a desk as a location proxy for “that wall” (figure 6.37), and then to where the sofas might “come out”, using a table (figure 6.38). After the usual difficulties with producing a point, Steve gestures with his mouse to the corner, his arm extending to touch the desk. When Arnie confirms “errrm yeh”, Steve then releases the mouse button and moves his mouse to a point on the floor to his right. He says “we could have the sofas coming (.) err (.) sort of out”. His arm briefly appears at its short length, as the cursor is over the background feature of the floor. This brief appearance of his arm occasions Arnie’s “yeh that’s what I was thinking” utterance. However, Steve moves his mouse slightly to the right so that it rests over the table, and extends to touch it. He then moves his mouse back to the floor area in front of him, as he says “out (.) like that”. Arnie then replies “yeh” again. The conversation then moves to a humorous exchange regarding the unusual production of a “good point”.

Steve’s “good point thing” is indeed an intricate sweep of the hand along the floor, in a semi-circular shape. It shows not only the potential positions of the “sofas”, but also their orientations in the way they come “sortuv out … out (.) like that” in a radial fashion, facing the TV. The table is used as a proximal object to the referent which will allow the system to extend the arm to an adequately correct length in order to produce the intricate pointing gesture.

Unfortunately, as with the first example, the possibility exists that participants might constitute the trajectory of the pointing arm as a gesture towards a different conversational referent. It is therefore interesting to note the work that is done in maintaining the identity of a referent when the structuring of objects in the system means that coincident reference in gesture and talk becomes a socially organised accomplishment. In other words, there is considerable work in referencing features of the virtual environment that both uses and takes into account the system resources provided.
6.8.4 Navigating based upon objects invoked by others

The fact that the system does not identify interactionally constituted objects as targets of gestures means that further problems arise when navigating by the targets of a co-participant’s actions. One consequence of a system definition of ‘objects’ at odds with interactionally constituted features of the environment is that what one user invokes for another may not be what the system provides for the other as a potential target for navigation. In the following example, a fragment of talk that precedes that in Example 6.19, Steve and Arnie are working out where the TV might be positioned.

Example 6.23 (Perspective: Steve)

Steve: in the corner (0.4)
Arnie: in- by- yeh (.) in the corner over there
Steve: how long (0.4) how long are our leads though (0.7)
Arnie: um (0.6) I dont think we know (0.5) but you don’t want leads trailing all over the place do you basically theres two- theres two possible places f- th- er for the TV (.) in that corner or in the opposite corner (.) the diag- diagonally opposite one=
Steve: =what about um right hang on I'm gonna point (0.8) we could (0.6) sthat pointing? (0.3) hang on I'm gonna turn (1.0) err right if you look (0.4) not towards me
Arnie: ((laughing)) oh my go- I've looked at you
The first part of the example involves a lengthy discussion regarding the “possible places … for the TV”. Steve then begins initiate the now-familiar preamble regarding a pointing gesture, “hang on I’m gonna point”. As discussed in section 6.8.2, Steve displays difficulties in seeing his pointing arm. However, of interest to us here is that
as he utters “err right, if you look-”, Arnie presses his ‘Other’ button, which his mouse is currently hovering over (circled in figure 6.40). At the same time, Steve has his (unextended) arm pointing along the wall that he is standing adjacent to. Arnie’s avatar is in front of him but facing away from him, with the corner off to Arnie’s right. If the object of Steve’s interactionally produced gesture were somehow correctly derived by the system then Arnie’s avatar would be rotated to view the correct region. Unfortunately, however, the navigation target for Arnie is identified as Steve’s avatar – because there is no system-defined object being pointed to. As a result, Arnie’s avatar ends up facing Steve’s at close proximity, with the result that Arnie’s avatar is rotationally further from viewing the relevant corner (figure 6.41).

Collaboratively constituting material features of the virtual environment has been shown to be no simple matter. Furniture World II differs from Furniture World I not in the structure of the environment, but in the use of explicit mechanisms for using that structure. This use was introduced in order to enable awareness between participants of each others’ object-focused actions. However, this example shows that target-based navigation takes this problem one step further. In the previous section it was seen that users have to explicitly work to establish an extended reference. However, this example shows that, on top of this, their co-participants may have real problems navigating to the target of those references. This is particularly true where system-defined objects are not adequate enough for use as proxies.

6.8.5 Difficulties in invoking objects as relevant in Furniture World II

To summarise this section, perhaps it is worth looking at one more example which clearly illustrates many of the problems with object definition that have been described. In this final example, 6.24, it is clear that the Furniture World system defines an object being worked with as very different from the objects constituted in the participants’ interaction. Tom is still (see Example 6.10) experiencing extremely slow movement, due to his machine being loaded down with additional unrelated processing from other networked users. As we join the conversation, Tom and Dave
are making plans to go and retrieve the lamp, which has been accidentally moved outside the room through one of the walls.

Example 6.24 (Perspective: Tom)

Tom: whereabouts is the lamp I’ll go see if I can get it
Dave: er: (2.0) jus in the wall there (1.0) see where I’m pointing?
Tom: (0.8) near the door sort of thing
Dave: no no in between the corner and the door
Tom: oh right
Dave: can you see where I’m pointing now?
Tom: ye- (0.3) I’m just trying to spin round but the things going so slowly its really annoying to try and spin round
Dave: if you click what I’m doing its pointing at the lamp now
Tom: (1.5) I’m holding that and its looking at you
Dave: doh (0.4) well its just outside the room where I’m pointing
Tom: oh there you go its got to you now (0.6) its got to where youre pointing

31 This and the previous utterance provide another example of the ‘subjective lenses’ phenomenon outlined in section 6.7.1.
Figure 6.42 – Dave’s view as he says, “Just click what I’m doing”

Figure 6.43 – Dave’s view as Tom says, “I’m holding that and its looking at you”

Figure 6.44 – Dave’s view as Tom says, “oh there you go … its got to where you’re pointing”
The speed of his computer means that Tom’s ability to quickly spin his view round is even more limited than usual. Dave says “click what I’m doing its pointing at the lamp now” (figure 6.42). However, when Tom clicks the button, his view centres on Dave’s avatar, rather than the target of Dave’s gesture (figure 6.43). This is due to Dave’s mouse having moved slightly from ‘over’ the lamp to ‘just off it’ (although he has no sense of this, because he cannot see the lamp behind the wall, and his arm remains the same length). The CVE now ‘sees’ Dave as pointing at a ‘background feature’ rather than an object. Thus, when Tom presses his ‘Other’ button, his view is positioned accordingly, centring on Dave’s avatar. As Dave begins to instead say that the lamp is “just outside the room where I’m pointing”, he moves his mouse very slightly. The cursor happens to reside back over the lamp again. Luckily, Tom still has his ‘Other’ button depressed, and his view now changes to orient to the lamp behind the wall to which the system now ‘believes’ Dave’s arm is pointing (figure 6.44). Tom then says, “oh there you go … its got to where you’re pointing”.

“Where I’m pointing” as a relevant feature of the virtual environment is mutually consistent to Dave and Tom, defined and re-defined through each action they produce. However, this point is done through the resources of the system which draws distinctions between background features and system objects. In this example, Dave is consistent in his definition of the location of the lamp. However, the CVE system is not consistent in its treatment of ‘what Dave is doing’. Small changes in Dave’s mouse movements cause vast changes to the participants’ views – and to their constitution of the location of the object.

It should be noted that this section is not highlighting that differences occur between talk and embodied work with objects. Where the system switches between the adequacy of the point as being at an object or not is ‘interactionally inconsistent’. Rather, in the cases of pointing and target-based navigating, going from ‘what object relevant to actions the system identifies’ and ‘what object relevant to actions the participants constitute’ forces an enormous amount of work on those participants.
Starting from the resources that the system provides, participants must accomplish shared points of reference through a moment-by-moment negotiation. Thus, it is not so much that the system imposes structure on the participants, but rather that the use of that structuring as a resource in work makes that work very difficult. Whether system resources are ‘tolerably correct’ in their definition of objects, or whether more work must be done to secure mutual orientation, is a cooperatively achieved task common in Furniture World II.

6.9 Summary and discussion

This targeted evaluation has described in detail the work done by participants in the Furniture World II trials. In sections 6.4, 6.5 and 6.6 the three interface implementations that were used to improve participants’ ability to organise their work with virtual objects were looked at. In sections 6.7 and 6.8, two consequences of these techniques have been described, which highlight some points about the structuring of CVE technologies. For clarity, each of these five issues will be summarised in this section.

6.9.1 Peripheral Lenses

Participants in the Furniture World II trials use peripheral lenses in improving their ability to see objects in the peripheral environment. Furthermore actions, that might have not been visible for a participant in Furniture World I, may be available in a distorted condition in peripheral lenses. It becomes difficult to distinguish participants’ orientations to peripheral lenses from other interface techniques. For example, extended actions or target-based navigation are often also used to view visual aspects of action in undistorted conditions. Further features were highlighted by another study, (Fraser et al., 1999), in which peripheral lenses and peripheral glancing were key to accomplishing interaction in a CVE (but which did not include target-based navigation functionality).
6.9.2 Extended actions

The three actions extended from the original Furniture World interface are grasping, pointing and looking. Each is used in a manner suggesting that intricate co-ordination of these actions is possible. The grasping representation is used as a resource in task discussion (rather than an explicit topic); the pointing representation is, on occasion, used to produce intricate referential gestures; and the view frustum is a key basis for designing actions such that others can see them.

6.9.3 Target-based navigation

In the cases where the target-based navigation algorithm determines the target as the object constituted in the participants’ interaction, it is a useful tool for overcoming the slow movement of gaze direction.

6.9.4 Subjective views in object-focused interaction

Two interface techniques – peripheral lenses and the extended view frustum – can pose problems for participants when organising their actions. This is due to an inability to see what another can see in the environment. The use of the peripheral lens technique by multiple participants means that they can view the same part of the environment at different levels of distortion. This may cause spatial relationships within the environment to be seen by each participant differently, even where they are viewing the same area. Visible view frusta not only show what another can see, but also intrude upon the environment, particularly where the frustum bisects a user’s view from the features of the environment they are working with. Moreover, the user whose view is denoted cannot see the effect that their position and orientation are having on the environment of their co-participant. Thus, extending views and extending actions can, in different ways, increase the difficulties for participants in constituting a shared environment.
6.9.5 Structuring objects in virtual interaction

The examples discussed show how objects as features of the virtual environment, are features of interaction that are invoked and tacitly negotiated as an on-going accomplishment. However, the CVE system categorises the structure of the world a priori, programmatically deciding the intricacy (what is an object, and what is just a feature of an object) and the validity (what is an object, and what is not) of particular interactionally constituted features of the environment.

Whilst these distinctions are also made in Furniture World I, they become increasingly problematic for participants interacting through the Furniture World II interface. The representations and techniques are more explicit than the original interface in their reliance upon definitions of objects – for example stretching to touch them, or turning to look at them. With this explicit notion, comes an increasing disparity between the object of the CVE system, and the object constituted in participant’s interaction.

6.10 Conclusions

This chapter has evaluated the redesigned interface produced for supporting object-focused interaction. It looked at the effects and use of the techniques employed. Furthermore, it looked at unanticipated issues that arise as a result of using those resources. These issues suggest a number of key points:

- Individual designs may be enhanced in certain ways.

- A toolkit for object-focused interaction in CVEs may draw upon the designs defined in chapter 5, but must be tailored to the unanticipated phenomena that arise in Furniture World II.

- The ways in which CVE systems structure their environment will affect and transform participants’ collaborative work with objects, and therefore that designers of virtual environments must carefully consider the ways that they
might approximate how collaborative users will constitute the environment and its structure.

- Interface resources are only as relevant as their use – in other words, neither ‘forcing users to learn to do things in the right way’ or ‘redesigning for users’ will accomplish ‘universal usability’ – but rather, that the ongoing accomplishment of distributed collaborative work will depend upon the relationship between design and use.

These issues will be further discussed in the following chapter, which examines how this analysis might be used to realise the contributions and findings of this thesis.