H. Gellersen et al. (eds.), ECSCW 2005: Proceedings of the Ninth European Conference on Computer-Supported Cooperative Work, 18-22 September 2005, Paris, France, 267–286.

Representations Can be Good Enough

Jacki O'Neill, Stefania Castellani, Antonietta Grasso, Frederic Roulland, and Peter Tolmie
Xerox Research Centre Europe, Grenoble, France
{name.surname}@xrce.xerox.com

Abstract. When working remotely with physical objects obvious problems of reference arise because of the lack of a mutually shared object. Systems aiming to support such work tend to be based on understandings of face-to-face interaction and frequently use video. However, video introduces new interactional problems. This paper describes a field study of remote interaction around objects that is telephone-centred, namely in a call centre for troubleshooting office devices. We describe how breakdowns in mutual orientation stem from three main problematics: 1) The inadequate fidelity of operators' support resources; 2) The lack of mutual access to indicative resources; 3) operators' lack of direct access to customers' actions and orientation. From this analysis, we have developed a design proposal for supporting such work. Rather than using video, we propose that utilising a linked problem representation would address these problems. To this end we describe our proposal for a bidirectional remote visualisation of the troubleshooting problem.

Introduction

A recurrent area of research interest in CSCW relates to how remotely situated people can work together when this work involves physical objects in the local environment of one or more of the participants (Fussell *et al.*, 2000; Gutwin & Penner, 2002; Kraut *et al.*, 1996; Kuzuoka *et al.*, 1994; 2000). Such work produces a number of interesting issues, centering on how to make the object at the local site available in someway to remote sites such that remote and local participants can work with it. When remote interactions take place around such objects obvious problems arise from the fact that the object is not mutually shared. What are trivial matters of reference in face-to-face situations, such as mutual orientation, establishing mutual understanding of referents, pointing, gesturing, knowing what people are doing or have done become problematic when participants are remote.

A number of different systems have been developed in an attempt to make local artefacts available remotely such that local and remote workers can work collaboratively on them. The prime approach to this problem so far has been to take face-to-face interaction around physical objects as the basis for systems design. Such systems aim to recreate aspects of face-to-face interaction around remote objects and tend to use video as the medium for bringing local objects to the remote site (e.g. GestureMan (Kuzuoka at al, 2000) and the work of Fussell *et al.* (e.g. 2004)). However, a common problem with video is that not only does it fail to recreate the richness of face-to-face interaction it also introduces new interactional problems for its users, something made evident in the early work on media spaces (Heath & Luff, 1991 & 1992).

Our research has taken an alternative approach to design in these circumstances. Rather than treating face-to-face interaction as a starting point, we began by examining a situation in which remote interaction around objects already occurs. Our field of interest was telephone support for copier-repair where local users (customers) have a problem and remote technical support experts (operators) attempt to talk the customers through troubleshooting that problem. This study follows in a tradition of studying the telephone-mediated work of call centres, both in general (see for e.g. Whalen, 1995; Bowers & Martin, 2000) and, more specifically call centres for large office devices (e.g. Whalen & Vinkhuyzen, 2001; Whalen et al, 2002). We present here the findings from a field study of this domain, with an emphasis on the methods the interactants use to establish co-orientation to and co-ordination of action around a non-mutually-shared object. Studying object work as it is carried out in a call centre enables us to examine the minimal support required to make such interactions effective. Thus avoiding many of the difficulties relating to how one might extract from the manifold richness of face-to-face interactions just which features are necessary for remote work. This is not to say that studying face-to-face interactions cannot give insights into how such interactions might be supported remotely, rather that the call centre provides an ideal opportunity to examine existing remote object-focused work from which new ideas for technology support might be derived. We now turn to, firstly, the existing work on supporting remote collaboration and, secondly, Whalen's work on remote support in a document machine call centre.

There are a number of systems designed to support remote collaboration around locally situated objects using video and audio in an attempt to create mutual co-presence with the object. One such system is GestureMan (Kuzuoka, 2000), a mobile robot with wireless video communication and a laser pointer, designed to support gesture between remote participants. It was tested in a series of naturalistic experiments described by Luff (2003) which uncovered new interactional problems arising from its use. Participants lacked reciprocal views, making acting on objects in the local and remote environment difficult because they could not design their conduct to be sensible and recognisable to other people. Mediation of action through the robot makes the coordination of such action difficult and orienting to objects also required work by both parties. Participants talked directly *about* the orientation itself, to overcome such problems. Luff concluded that conduct and ecology are reflexively related and by creating new environments, with technology supporting remote participants, the relation between action and the relevant ecology may be fractured, causing interactional problems.

Kraut, Fussell, Siegel and others carried out a series of experiments to understand the key visual information required in collaborative physical tasks (their tasks being bicycle repair and robot building) and thus the requirements of technologies to support such tasks (e.g. Kraut, Fussell & Siegal, 2003; Fussell, Setlock & Parker, 2003; Fussell *et al.*, 2004). They described how visual information is used to time instructions and for pointing and other deictic expressions. They implemented a number of different video arrangements (see for example, Fussell, Setlock & Kraut, 2003). However, they too found that video introduced new interactional problems, including dislocation of gesture and lack of reciprocal views. They concluded that the task view was the most important and that gesture needed to be embedded in the task environment. To this end a new system, DOVE (Drawing Over Video Environment), was developed to enable gesture around the task object (Ou *et al*, 2003). DOVE enables participants to share a workspace via video, with representational gestures and pointing done by over-laying pen-based gestures on the video stream. Testing (Fussell *et al.*, 2004) showed that the system was primarily used for pointing (75% of drawings).

To summarise then, a number of systems have been developed to support remote work around physical objects which attempt to recreate the salient features of face-to-face interaction. However, in the creation of new environments for interaction many of these systems fragment the relationship between action and its relevant environment, introducing new interactional problems which can make even seemingly simple activities problematic (Luff, 2003).

In this paper we examine the features of object-focused work in an *already established* remote environment, that of a machine troubleshooting call centre. Whalen and Vinkhuyzen (2001) studied call centres in the same domain, describing how the expert system, implemented in the call centre to enable the operators (Customer Service and Support Representatives, CSSRs) to diagnose machine problems and direct calls to the relevant hardware or software support services, embodied misconceptions about knowledge and expertise. Primarily that the expertise could reside within the system, utilised by non-expert CSSRs and that the CSSR-customer interaction was a unproblematic one, requiring merely that the CSSR enters, exactly, the customers problem report and repeats verbatim questions from the system. However, in practice this ignores how the CSSRs working knowledge of the technology and sensitivity to the user's circumstances, by necessity, shape the way they handle the problem. Thus Whalen and Vinkhuyzen outlined the common-sense practices used by operators to circumvent the system and do practical troubleshooting with the remote party. In effect, because of the non-expert status of the CSSRs this troubleshooting tended to revolve around arriving at a point where a service call could legitimately be made. The CSSRs made judgments and interpretations of the customers input, but with little machine knowledge they had to primarily use their interactional understandings, orienting to the call as a service encounter. This is in contrast to the work of operators with machine knowledge examined in both Whalen and Vinkhuyzen study and our study, who can utilise these understandings along with their interactional understandings to diagnose and fix machine problems. In the Whalen and Vinkhuyzen call centre the expert operator was the exception, whereas in ours they were the rule. In the call centre described in this paper, the operators are trained and the expertise is seen to reside with them, with support from a knowledge base rather than residing within the (expert) system. Indeed, even their titles could be seen to reflect this, that is Technical Support as opposed to Customer Service and Support Representatives. Some features of the work however can be seen in common and where such similarities occur they will be highlighted. However, although both papers examine work in similar organisations, the organisational process of the call centres differs (the non-experts and expert system in one versus the trained staff and knowledge base in the other) as does the analytical perspective. Whereas Whalen and Vinkhuyzen describe in detail the 'expert system; CSSR; customer' interaction we are primarily examining the 'technical support; customer; machine' interaction.

For such interaction the critical requirement is the ability to mutually attend to the machine and engage with it, not necessarily to be actually co-present or to recreate co-presence with that object. One issue that this work brings to light, and which will be explored in the discussion, is the relationship between the work to be supported and the optimal nature of the support. We suggest that a representation of the troubleshooting problem, from herein called the 'problem representation' can be good enough to support such interaction. A proposal for how such a problem representation might be designed is given later on in this paper. Although it has yet to be implemented, the proposal is firmly grounded in this research into remote work with physical objects. The work of remote experts giving help to customers attempting to fix problems with their office devices will be described in the next section, followed by the proposed problem representation which will then be discussed in the light of the previous work described.

Fieldwork Observations

The field work consisted of a three week ethnographic study of a European Call Centre for a copier and office device company. The study involved observing the operators at work. Data was collected through field notes, video and audio recordings[1]. The call centre in question provides telephone support across Europe for customers with problems with their office devices (copiers, printers, MFDs, etc.). Operators lead the customer through a process of troubleshooting the problematic device. This work involves a number of activities:

- Operators first elicit an initial problem description from the customers. This initial problem description is often partial and the full description of the problem, as it appears to the customer, may be provided during the course of the interaction. For instance multiple symptoms will not necessarily be described all at once.
- Next operators and customers collaboratively work up the initial description into a fuller description from which they can begin to arrive at possible solutions. Often the operators require additional information about the machine, which they get via the customer. This may involve getting the customer to carry out tests on the machine. This collaborative production of the problem description was also noted in Whalen and Vinkhuyzen where with non-expert CSSRs it caused problems of diagnosis which they observed did not occur where the CSSR had the expertise to probe the customer further for a more precise description.
- Then the operators and customers work collaboratively to troubleshoot the machine, with operators giving the customers instructions to carry out and customers reporting back on the results of their actions.

In this paper we will show how operators and customers work together to create and maintain a mutual orientation to the device through talk. It is this shared orientation that enables the remote troubleshooting to take place. Operators have a number of methods for dealing with their lack of direct access to the machine in question and these will be examined, along with how and where breakdowns in this mutual orientation may occur and how such breakdowns are repaired.

Establishing Shared Referents

Operators and customers engage in interactional work to establish shared referents in the absence of mutual access to the device. We elaborate here on how the shared understandings that Whelan and Vinkhuyzen noted with their expert users are arrived at. An important aspect of this is how operators and customers question one another's descriptions to ensure they are referring to the same thing.

---

1 For legal reasons only the operator side of telephone conversations could be recorded on audio. Customer utterances were recorded in the field notes.

For example, in Extract 1, the customer reports a problem with a particular part of the machine, 'the paper feed'. The operator questions the customer, re-describing the referent according to its use ('where you put the originals in').

Extract 1[2]

1. C - I've got a problem with the paper feed
2. O: Um hum (.) You are talking sorry you are talking where you put the originals in aren't you

Operators also perform checks to ensure that the customer knows what part they are referring to. Thus one method of establishing shared referents is to reformulate descriptions according to different features of that referent, such as function, colour, shape, relative position, and so on.

Operators frequently use such descriptors to make their instructions understandable, adapting their utterances for the customer. Since many customers have relatively little technical knowledge about office printing devices, operators often use vernacular-type descriptions, occasionally with textual indicators (e.g. 'Can you just open the exit cover for me, the one that says CopierCo[3] on it'), to indicate parts rather than relying on technical terminology alone.

The manufacture of the machine with different coloured parts aids this location of referents. Operators know their machines well, describing machine parts from memory in such a way as to make it easy for the customer to locate them. In their work as operators they have evolved a comprehensive grammar of reference, reformulation and redirection. This stands in contrast to the 'helpers' seen in many of the previous studies outlined above who had little expertise in such remote help giving. Where the customer is able to locate the parts easily and follow the operator's instructions it is not necessary for the operator to be able to see what the customer is doing or where the customer is looking. The customer's verbal responses, combined with the operator's knowledge of the machine, are often enough for the operator to be able to indicate and clarify referents and give sequential instructions.

Directing Customers Through Sequences of Actions

As we have already indicated, operators must give instructions to customers regarding parts of the machine and/or sequences of actions to be carried out on those parts, even though they themselves do not have direct access to the machine. They therefore make use of the methods described above for accurately and adequately giving instructions to customers. Although operators frequently devised instructions 'off-the-top-of-their-heads', at times they utilised additional

---

[2] Where customer turns were not recorded in the field notes they are omitted.

[3] CopierCo is a fictitious name for purposes of anonymity.

resources to situate their instructions in relation to the machine. These resources are comprised of:

- *The knowledge base*: operators have access to a searchable knowledge base of solutions on their PCs, containing images of the various instructions. Operators use this as a visual aid from which instructions can be devised. For instance, one operator was observed pointing at an image on the knowledge base while instructing a customer through a set of actions, using colour and positioning descriptors to identify the parts:

Extract 2

1. O: ok and where you have door a you have like um a set of four grey rollers
2. <as she says this she points to them in the picture on her screen>[4]
3. on er a metal bar just above that there's a piece of black plastic and […]

- *Menu maps*: operators use menu maps to lead the customer through their on-screen options. We shall examine the adequacy of menu maps below.
- *Miming*: operators are frequently seen miming actions whilst simultaneously describing them to the customer. Whalen and Vinkhuyzens expert CSSR was also seen to gesture while talking. Miming is used in the absence of the device to establish the sequence of actions that the customer must undertake. As with the pointing above, operators frequently used gestures despite this resource being unavailable to the customer.
- *Going to the machine*: most machine models are available in the call centre. Operators often leave their desks and physically go to these devices 'to see what the customer is seeing', enabling them to describe parts and action sequences more precisely.

The above resources enable the operators, in the absence of direct access to the problematic device, to visualise the machine and the sequence of actions to be carried out upon it. Although these resources, along with an operator's knowledge of the machine, are often adequate for troubleshooting, there are two problems that can arise with their use. Firstly, these are *generic* resources *representing* the problem device, not the problem device *itself* and thus their fidelity is not always adequate for troubleshooting. Secondly, the indicative information involved is not available to the customer, making it a lost resource and requiring the operator to translate it into verbal instructions.

These issues can feed into situations where the troubleshooting process encounters trouble or even breaks down completely. Such trouble arises for the operator in establishing what it is that is going on at the customer end and for the customer in attempting to put the operator's instructions into practice.

---

[4] Text in < > brackets indicates an action.

Establishing the State of the Machine and Related Artefacts

As suggested above, at times it is not enough for the purposes of troubleshooting to know a machine in general, rather the specific state of *this machine here* or its related artefacts, such as copies, becomes important. At various points in the interaction, operators need to establish what the state of the machine or related artefacts is. Their understanding is of necessity mediated by the customer.

Operators may check the state of the machine to enable them to give relevant and appropriate instructions. For example, they may ask if all the doors are closed. Operators also ask customers to tell them what some part or other of the machine looks like because knowing what a machine in general looks like is not the same as knowing what the machine looks like in just this instance. Yet it is often features of this particular machine *here and now* which are pertinent for troubleshooting. To uncover the relevant features here and now (or, as Garfinkel would put it, the *haecceities* of the problem (Garfinkel & Weider, 1992)), operators get the customers to examine their machine or to elaborate on prior descriptions.

Extract 3

1 O: Um and can you tell me when you look in is the tray still lying flat or is it
2 er a bit off does it look as though it's skewed by any chance?
3 O: It does look? Ok I just wonder if we can sort that out now

In Extract 3 the operator asks for information on the tray, proposing possible alternative scenarios, 'lying flat' or 'skewed' (1-2), to help the customer understand what they are looking for. These alternatives relate to the possible causes of the problem and thus are relevant for the troubleshooting process. The customer confirms that it seems to be skewed and the operator begins the process of rectifying the problem (3). By asking the customer about the state of the tray the operator is exploring ways of narrowing down the problem space (either by eliminating or finding a cause of trouble if the tray is flat or skewed, respectively).

Another method used is 'drilling down', where the operator asks a series of successive questions to get all the necessary detail and ensure a common understanding. For example, refining an understanding of an image quality problem by questioning the customer on the state of the copy, e.g. 'Is it all creased up?', 'Is the whole page creased up or half the page?', and so on. Both of these methods can help operators to refine the problem space, propose causes and suggest solutions. However the lack of direct access can result in incorrect instructions, for example asking the customer to 'open up the top cover' when it is already open or directing the customer to 'a blue plastic guide' when it is in fact green. Customers, of course, are able to and do correct such mistakes.

Where Mutual Orientation Breaks Down

Where customers cannot identify the part or other referent which the operator is describing, the operator must attempt to disambiguate the referent. This has to happen with little knowledge of the customer's actual orientation and the state of the machine. Two examples of this are presented below, the first in which the customer cannot locate a part, and the second in which the customer cannot find an entry on a menu map.

Disambiguating confusion: locating parts

Difficulties can arise in locating and identifying physical parts of the machine and the only methods available to the parties to resolve these involve further talk. This might include repeating instructions, reformulating descriptors and terms, or elaborating descriptions (e.g. by describing relative position or functional features (where the paper comes in/goes out, and so on)). Several examples were seen where understanding relative directions, in particular, right and left caused considerable trouble for the customer and took much effort to resolve. Extract 4 shows the work to resolve the location of some doors.

Extract 4

1 O: ok it's probably saying open the upper left hand side door? Probably one of right
2 there's two doors there that you open there's the first door that opens downwards and
3 then there's a door in front of that which is the hot area of the machine so you don't
4 touch that area and you just need to check that to see if there's any paper sticking
5 out that you can actually remove just to see if if you know you can remove it there
6 O: yeah course no problem take your time
7 <C goes away> (long wait) <C returns>
8 O: hello
9 C – can only find the big door and the little side door.
10 O: Yeah yeah yeah you know the when you slide the finisher away from the machine
11 you can open the upper left hand side door that opens downwards? Then just in front
12 of that there's another door and that's where the hot area of the machine is so don't
13 you don't touch the roller or anything just jus you're just looking for any paper that
14 you're able to actually
15 C – can't see any paper and there's only one door
16 O: No the the with the first bit you've got an upper left-hand side door and a lower
17 left-hand side door now the upper left-hand side door has two doors that you can open
18 the lower one (doesn't)
19 C – I'll go and check.
20 O: OK no problem
21 <C goes away> (long wait) <C returns>
22 C – I managed to retrieve the paper

In lines 1-5 the operator gives a detailed description of what the customer should do, including describing the doors to open according to their relative locations and opening mechanisms. The customer goes to do this, but returns unable to locate the right doors (9). The operator reconfigures her description twice (10-14 and 16-18) the second time because of the customers contradiction 'there's only one door' (15). This time the customer succeeds in locating the door and retrieving the paper (22).

We can see that the work in this case arose because the customer could not find what the operator was referring to. This was problematic to resolve because the customer only had limited understanding of what exactly the operator needed her to orient to. The only methods available to the operator and customer to resolve these issues and disambiguate the instructions are those available through further talk. Where instructions do not seem to be working, operators reiterate and reconfigure their descriptions, often repeatedly, both checking that the customer is doing the right thing, and reformulating them to make them more understandable. However, if operators had a better understanding of exactly what customers were orienting to, such reformulation would be far more straightforward. Additionally if customers had easy access to what the operator was referencing instruction would be more straightforward. The operator also does not know exactly what it is that the customer is doing at any one point, so cannot help the customer by correcting his errors as he makes them.

Disambiguating confusion: menu maps

Extract 5 is a further example where the customer cannot locate what the operator is directing him to, this time with regard to on-machine menus.

Extract 5

O: Ok can I get you to go into the front panel and select menus
O: Ok then scroll until you see printer set up menu
C – printer set up menu
O: Ok then scroll until you see energy star/power saver
C – energy star
(silence)
O: you're not seeing it no
<C reads list of menu options >
O: ok so I'm just quickly going through the menu map that I have here myself ok
<O looks at the menu map on screen> <C reads list of menu options again quietly>
O: ok can I just bear put you on hold for one second just want to check something with a colleague of mine
<O talks to colleague who tells him that if it's not there it is turned off. Returns to desk>
O: hello karl?
O: Yeah sorry about that delay there ok yeah if that's not appear if that option's not appearing on your front panel then it it would mean that that option has already been selected it has been switched off already so that it's basically it won't say after an hour or something go into this standby mode () it's
C – it does go into standby
O: does sorry
O: ok see if if it is going into that that option should be
O: ok
O: ok because that's what I've been advised if it's not showing up on on that menu page then it has been disabled in the machine. […]

Initially the customer is following the operator's instructions without problem (1-3), with the customer repeating back what the operator has directed him to (3). Then the operator asks the customer to scroll to the energy star (4) which the customer repeats as before (5). Trouble is signalled by a silence (6) which prompts the operator to propose an explanation for the silence 'you're not seeing it no' (7). The customer reads through the list of menu options (8), thereby making them available to the operator[5]. The operator checks his on-screen menu map. Operators often use menu maps as a way of visualising what a customer is seeing and thus enabling them to direct the customer through a series of actions. Menu maps are a stand-in for the fact that the device on which they are working is not mutually shared. However, as menu maps are an idealised instance of the menu, they *do not show what the customer can actually see*. In most cases, of course, this may be good enough. Whilst the exact labelling of menu options may not always be the same as on a customer's machine, for the most part operators can easily get round this by saying 'can you see something like…' or offering several different variations on likely names[6]. However in this case, the fact that the menu map is not the same as the customer's actual menu is more tricky to deal with.

While the operator is checking his menu map the customer re-reads the options from his own more quietly (10). Having checked the menu map and finding the energy star on it as expected, the operator excuses himself (11-12) and goes to check with a colleague (13), who explains that if it is not there then it has already been switched off. The operator explains this to the customer (15-18) but the customer disputes it (19 and between turns 20 & 21, 21 & 22 and 22 & 23[7]). The operator responds by reiterating that it has been disabled (21 & 23-24), then moves on[8]. Thus this difference between the idealised version of the menu map and the customer's actual menu required additional work to 'resolve'[9], with the operator first trying his own resources then having to take time out of the call to consult with a colleague.

We can see, then, that such trouble arises where what the customer can see appears to differ from the operator's description or where the customer just cannot see, for whatever reason, what it is the operator is describing. There is a difference

---

5 One of the features of the phone is that just what is and is not shared is readily available to both parties, as demonstrated here.

6 Although this is a noted problem in the non-English language groups if the operators are using English menu-maps as their translations can be quite different from the formulised menu-map translations.

7 Customers wording between these turns not available.

8 Interestingly this non-acceptance of the customers assertion that the machine does go into standby resembles somewhat Martin & Rouncefield's (2003) finding that only where the bank actually has a letter sent by the customer are they accountable for it, if they only have records they are only accountable to them and if they have no records of the object it does not exist. Here of course it is the behavior of the object rather than the object itself that is being held up to question, but it seems that where the customers report of the behavior of the object differs from some expected behavior, as confirmed by a colleague, it is the customers account that can be disregarded.

9 Indeed, we can only say it was resolved in that the operator moved on to other troubleshooting activities, rather than that consensus was reached between the customer and the operator.

between locating physical parts and locating menu items which stems from a level of certainty. That is, the operator can be fairly certain that a part of the machine, doors, handles, etc., will be there for a particular model of the machine. In that case if the customer cannot locate it, it makes sense to reformulate and reiterate the directions until the customer *can*. In such situations it is assumed that the source of the trouble lies in the direction of the customer. However, with menus it is a rather different situation. Menus can be reconfigured in a way that changes them but which is not necessarily obvious to the user or presumptively certain to an operator. For instance, in the above example one possibility is that the energy saving feature had been switched off. Hence the energy star was no longer present on the menu, although this was disputed by the customer. Hardware can of course be reconfigured, but the presence or absence of a finisher, for example, is relatively easy to determine. Changing the settings on menus, by contrast, can effect what does or does not appear. Thus an operator can be less certain of the source of trouble when such issues arise .

Instruction in practice

Giving and following instructions is a collaborative activity designed for and by the co-participants. Instructions are designed to be timely and appropriate. Operators attempt to fit the instructions with customer activities and their situation (e.g. step-by-step if at the machine, in bigger chunks if having to move between the machine and the phone) and use appropriate language. As in Extract 5 (1-5) operators often time their instructions according to the activities of the customer. However, as described above operators only have limited access to what the customer is doing and orienting to. Access is limited to what is provided through customer feedback, though operators do, of course, work using assumptions of what is happening on the basis of their understandings of how such troubleshooting episodes usually proceed. However, as shown in Extract 6, this presumption is not always equal to overcoming the absence of personal access.

Extract 6

1 O: That's where the paper would normally um feed through ok so er it's just in there
2 that you're feeding the paper that you're putting the page in?
3 O: Is it?
4 O: Hello?
5 C – yes

In this sequence, which arose during a call where the customer was having problems following the operator's instructions, the operator asked the customer a question (1-2) then, on receiving no reply, twice prompted the customer for an answer (3-4). This occurred because the operator did not know what was going on at the customer site. Indeed, the operator remarked at the time that 'sometimes you wonder what they are doing'. Also later in the call the operator repeatedly asked the customer questions along the lines of 'Does that make sense?' 'Is that working?'. This is because when they get no feedback from the customer they have to try to work out what is going on. The production of such utterances in the absence of feedback is a systematic feature of talk (Sacks, 1992). The absence of a response to a question is highly accountable and typically leads to truncated repetitions such as the one visible in Extract 6 (line 3) (Atkinson & Drew, 1979; Heritage, 1984; Schegloff, 1972)[10] . The difficulty with this call was compounded by the fact that the customer had to put down the phone in order to follow the operator's advice.

Currently, then, both the customer and the operator must work together to disambiguate referents and instructions and to establish a mutual orientation to the object. Although this often works well, it can create difficulties where parts cannot be identified or instructions followed. Where this happens the mutual orientation to the object is lost.

To summarise, troubleshooting the machine is a collaborative activity and is based on a mutual orientation to the device. However, the lack of mutual access to the problem device can result in breakdowns in this mutual orientation which stem from a number of problematics:

1) The operator's resources to visualise the problem device are *generic* resources representing some type of device in general rather than the haecceities of this particular problem. Consequently, their fidelity is not always adequate for troubleshooting.
2) The lack of mutual access to indicative resources means that the operator's gestures are not available to the customer. Instead they require translation through talk. Similarly, customers can only indicate the source of their misunderstanding through talk.
3) The customer's orientation and actions are not directly available to the operator. In that case the operator must rely on a customer's feedback to situate and disambiguate instructions.

In the next section we will outline a design proposal to address these problematics.

Bidirectional Visualisation of the Troubleshooting Problem

To address the problematics outlined above (generic rather than indexical resources, lack of mutual access to indicative resources, lack of direct access to customers orientation and actions) we examined ways in which the features of the actual troubled device itself might be made available to both parties. Primary here is finding ways to enable them to mutually orient to it, share indicative

10 The conditional relevance of utterances in these kinds of situations is more generally discussed by Schegloff (1968).

information such as gesture, and enable customer actions to become available to the operator. One such way is to provide the interacting parties with a *representation of the troubleshooting problem itself.* Such a representation would provide a resource for both coming to an understanding of the problem and mutual orientation and interaction. To this end, support could come from providing a shared object (i.e. the problem representation) to which customers and operators could mutually orient and refer which would reflect the actions of the customer. One of the crucial aspects of the telephone is that it gives a clear understanding for both parties of what does and does not fall within the shared space. Therefore, any solution that will be supporting this interaction should do the same, rather than creating the additional problems which arise where the boundaries and extent of the shared space as against the local space are not clear. To achieve this we propose to use a distinct representation of the machine and its troubles rather than focusing on video to connect the local and remote parties. Using this kind of representation offers a number of potential advantages. These include a ready recognisability of what is shared or purely local and low overheads in equipment.

Making Use of Representations

The design proposal outlined here is based around the creation of a bi-directional, shared visualisation of the troubleshooting problem (BDV). This problem representation will consist of a linked 3-D model of the device and a number of means of interacting with this model. The BDV will be presented on the device itself at the local site (on the kinds of medium sized screens increasingly available with modern devices) and on the technical support operator's terminal at the remote site. The representation is *linked* to the device itself, such that actions on the device are shown on the representation, e.g. if a user opens a door, that door will appear open on the representation. This is enabled through the many sensors that already reside on such devices. In addition both the customer and operator are able to indicate parts on the machine, and the operator is able to demonstrate visually actions which should be performed (for example, lifting a handle and sliding a toner cartridge out of the machine). The customer will access technical support through audio-visual communication channels located on the machine itself. The audio channel will enable the customer to converse with the operator. The visual channel will show the BDV. Thus the machine becomes the infrastructural mediator between users and technical support.

The BDV enables *both parties* to have a real time understanding of the actions which are being or should be performed on the machine. These provide a resource for overcoming the troubleshooting problems we have described. The machine will enhance an operator's understanding of the problem and thus aid the discovery of a solution. It will then mediate between the operator and the customer enabling them to mutually arrive at a solution despite not having mutual access to the problem source, i.e. the machine. The solution we envisage will allow:

- The creation of a dynamic virtual visual representation of the troubleshooting problem including a visual representation of the machine and interaction controls.
- Customers and remote operators to access a personalised view of the representation, which they can manipulate in a coordinated way where interactions on one side are captured, transmitted and appropriately made visible to the other side.
- Customers and operators to identify in their representation a component of the machine by indicating it.
- Operators to define on their view of the representation actions to be performed by users on selected components of the machine. This will be achieved through visual images, animations and descriptions of solutions being dragged and dropped from other resources, for instance the knowledge base. These actions will be transmitted to the customer via their representation.
- Customers to interact with the representation by manipulating the machine itself. That is sensed actions that the customer carries out on the machine will be shown on the representations. These actions will be transmitted to the operator via their representation.
- The local device menus to be made available to the remote operator so that they can direct the user through the correct navigation path. This is particularly pitched at problems and solutions which involve the user navigating the menu (as was shown, for example, in Extract 5).

A number of benefits could arise from using the BDV. We propose that these would both give advantages over the current situation and provide support for the troubleshooting interaction that is *good enough* at minimum interactional and equipment overheads. Benefits include the fact that many aspects of the state of the machine, such as doors open, trays pulled out, etc. would be evident to the operators without having to ask the customers. In addition changes to the state of the machine would enable the operators to get an understanding of the customer's actions, that is as the customer opened doors, removed machine parts and so on this would be represented on the operators BDV enabling them to 'see' what the customer was doing. Operators would be able to indicate parts and actions to the customer and customers would be able to indicate parts to the operators. Situating the instructions in the stream of activity would be aided by the representation as the operator would be able to 'see' what the customer had done more or less as it happened and thus give the next instruction. Reciprocal viewpoints are supported and operators and customers should be able to co-ordinate and co-orient around the representation of the object. Although just as with any other tool or artefact to be used during the service interaction, the BDV would have to be weaved into the interaction with the customer (Whalen, Whalen & Henderson, 2002). This is of course is already the case with the operators existing tools and indeed the machine itself (as we could see with the work to situate instructions within the stream of activity on the device). To this end the design of the BDV will need to take into account its use in interaction, and how exactly the features described will be implemented will need to be specified during the design process. In addition, like other such tools it is likely to introduce its own specific interactional difficulties which will only come to light upon implementation. However, with careful and iterative design we feel that the BDV could offer a useful alternative, in appropriate domains, to previously specified solutions and facilitate current interaction. The solution also contains non-representational aspects for menu-based instructions, where the operator can view the same interface as the customer.

The BDV is not designed to be an expert or other such system, rather it is a *communication tool* to be used by technical support and the customer alongside the audio interaction in troubleshooting the device. Although the representation is an idealised version of the object rather than the object itself, because it is tied to the actual machine, it is closer to the object than those representations already used (menu maps, machines, etc.). Although at first it may seem to be a relatively basic and simple representation, this seemingly shallow representation is actually able to capture salient indexical information so that the haecceities, the 'just thisness' of the problem (Garfinkel and Weider, 1992) can be explored and revealed. Here the focus is solely upon making available the orders of detail relevant to getting the troubleshooting job done instead of leaving the interactants still in need of uncovering saliency from a relatively undifferentiated video stream.

Discussion

The field work exposed three areas of work that operators and customers do to make remote troubleshooting work. These are:

1) *Establishing shared referents and mutual orientation* to the device through talk. For example, operators question customers, reformulate descriptions, use appropriate non-technical language, and so on.
2) *Establishing the state of the machine*. Customers mediate between the machine and technical support. Operators use checks, drilling down, offering proposals, etc. to narrow down the problem space. Customers report back on actions they have performed and resultant machine status.
3) *Situating instructions*. Operators are knowledgeable about the machines and have additional resources for visualising the device, thus supporting their

interaction with the user despite not having direct access to the device or the user's orientation or actions upon it.

Breakdowns in mutual orientation stem from three main problematics:

1) The inadequate fidelity of operators' generic support resources.
2) The lack of mutual access to indicative resources.
3) The operators' lack of direct access to customers' actions and orientation.

The BDV aims to address these problems by providing a shared object around which both parties can mutually orient. This shared object, being a representation of the device, is linked to the device itself thereby simultaneously increasing fidelity to the specific problem device and enabling the operator to view many of the customer's actions as they are undertaken. In addition indicative resources are provided to both parties.

The related work we described earlier, in particular that by Fussell et al, has tended to examine the visual information available in face-to-face interaction in a decompositional way. For example Fussell, Kraut & Siegal (2000) describe how different types of the visual information (from gaze to participants bodies and actions) can be used in conversational grounding. This leads them to video as a mechanism for recreating this shared visual space. Rather than focusing on visual information in a decompositional way, examining the detailed practices of those engaged in remote object-focused work shows how talk, referring, etc. is embedded in the circumstances of getting the work done. By focusing on the practical work of troubleshooting we have begun to get an understanding of where troubles occur in this work. We have seen that although customers and operators are often able to establish a mutual orientation to the device in question, asymmetries of access can also all too easily result in their mutual orientation breaking down. Where it does break down the parties have to rely on further talk to re-establish it and this can involve considerably more interactional work. This focus on the work to establish a mutual orientation to the non-mutually shared device led, in this case, to the idea that a problem representation could be good enough to support the troubleshooting interaction.

Another difference between our own study and the others that we have discussed is that we have studied the work of expert givers of help: people who are trained and work in the context of providing remote help on a day to day basis. Other approaches have resorted to using 'subjects' with no particular experience in help giving, relying instead upon the articulation of provided instructions. Our experts were seen to have developed skills in remote help giving, from hiding or accounting for the use of the system during interaction to miming the actions as they describe them to the customer. For example, in the bicycle repair task described in Kraut, Fussell & Siegal (2003), one issue was that helpers did not know when to intervene, yet here we can see that intervening as such (deciding when help should be provided) is not generally a problem. There are some difficulties with situating instructions but the operators have developed a grammar for appropriate instruction and reformulation.

The problem representation described here has a number of features which lead us to propose that it will be 'good enough' to accomplish this required order of enhancement. These features include: the ways in which parties could mutually orient their activities around the representation; an appropriate level of indicative resources; the fact that the state of the machine is available to both local and remote parties; and the linkage of the resource across sites allows for a high degree of fidelity to the troubleshooting problem itself. In addition, our solution is not about removing expertise from the hands of technical support, rather it is a communication tool designed to enable them to apply their knowledge more easily. That is, by providing some access to the remote object. Whalen and Vinkhuyzen comment on how CSSRs are disadvantaged by their lack of access to the remote object i.e. the machine and it's artefacts such as print outs, but that they have the rich resource of natural language to help them get an understanding of the machine problem but not the expertise to use it. In contrast Technical Support do have the expertise, however there are still aspects of the Technical Support-Customer interaction that can prove troublesome because of non-mutually shared access to the device. The BDV is an attempt to address this.

There are also a number of reasons which suggest that problem representations are more suitable than video for supporting *this* work. In particular these relate to the ways in which shared representations of this order should avoid the problems of fractured ecologies that video based systems introduce, since a common understanding of reciprocal views should be easy to achieve. In addition, the system outlined here is a more economical arrangement for the task at hand. Customers want to spend minimum effort troubleshooting their machines, so it needs to be made as simple and effective as possible. It is a solution that has the minimum overhead for all the parties concerned and is based on existing device features: sensors, medium-sized screens and high quality GUI (found on newer devices). Our solution does not require the user to wear or have any special equipment. In this situation low-cost video is certainly not likely to be good enough for many of the actual problems, whereas a good problem representation can have the advantage of clarity by not relying upon camera angles and orientations. Furthermore, it seems likely that if static cameras were used the number of cameras required would be prohibitive whilst the use of a mobile camera would negate many of the proposed benefits, requiring the operator to direct the customer to move the camera to the appropriate areas of the machine. Thus considering the limitations of other support and considering the actual requirements of the task, a problem representation tied to the actual object is likely to be 'good enough' for many of the kinds of troubleshooting that involve the participation of experts at remote sites.

Further research is still required, most particularly with regard to implementing and testing such a representation, but we hope that this paper has begun to contribute to an understanding of where other kinds of representation might be best suited to supporting interactions around remote objects. There are specific aspects of office devices that lend themselves to these kinds of representations well : newer models already have larger interfaces on them and they already have many sensors, allowing a degree of fidelity between the object and the representation that might be harder to accomplish in some domains. There are however other domains where similar levels of fidelity are available, another massive domain is vehicle repair. Vehicles are increasingly fitted with wireless technology and multiple sensors. The basic requirements of domains where such representations might be appropriate domains where mechanical manipulation of parts is required, there is the ability to repair on site and sensing infrastructure is viable. The size of many kinds of devices also makes low cost video solutions less appropriate than they would be for, say, a desktop task. It therefore seems sensible to suggest that design should be for the particular work-at-hand. That is, different work is likely to be more or less suited to different orders of representation.

Beyond all this, questions can also be posed regarding what can and should be represented and what adequate fidelity of a representation to an object might amount to in practice. For example, in the case outlined here will the proposed solution be adequate for specific aspects of the task such as instructing a customer through on-screen menus or for understanding and transferring information on the image quality of copies? These are issues that are subject to further investigation in the course of implementing the system described here.

So, to sum up, in this paper we have described the troubleshooting practices of remote experts and customers in order to delineate our reasons for proposing a different approach to designing support for such work. Other approaches have proposed video-based systems to recreate features of the face-to-face situation. Our research, by contrast, has suggested that, for many situations, a representation of the troubleshooting problem, tied to the source of the problem itself, would be ‘good enough’.

References

Atkinson, J M, and Drew, P (1979) *Order in Court: the Organization of Verbal Interaction in Judicial Settings*, London: Macmillan

Bowers, J., and Martin, D. (2000). Machinery In The New Factories: Talk and Technology in a Bank’s Call Centre. *Proceedings of CSCW 2000*, © 2000 ACM.

Fussell, S R, Kraut, R E, and Siegel, J (2000) ‘Coordination of communication: effects of shared visual context on collaborative work’, *Proceedings of the CSCW 2000 Conference on Computer Supported Cooperative Work*, 21-30, New York: ACM

Fussell, S., Setlock, L. & Kraut, R. (2003) Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. *CHI 2003*. 513-520

Fussell, S., Setlock, L. & Parker, E. (2003) Where do helpers look? Gaze targets during collaborative physical tasks. *CHI 2003 New Horizons*. 768-769

Fussell, S., Setlock, l., Yang, J., Ou, J., Mauer, E. & Kramer, A. (2004) Gestures over video streams to support remote collaboration on physical tasks. *HCI. 19*. 273-309

Garfinkel, H and Wieder, D L (1992) 'Two incommensurable, asymmetrically alternate, technologies of social analysis', in G Watson and R M Seiler (eds) *Talk in context: contributions to ethnomethodology*, New York: Sage, 175-206

Gutwin, C and Penner, R (2002) 'Improving interpretation of remote gestures with telepointer traces', *Proceedings of the CSCW 2002 Conference on Computer Supported Cooperative Work*, 49-57, New York: ACM

Heath, C & Luff, P (1991) Disembodied conduct: communication through video in a multimedia office environment. *Proc. Of CHI'91*. ACM Press. 99-103

Heath, C. & Luff, P. (1992) Media Spaces and communicative Asymmetries: Preliminary observations of Video-mediated Interaction. *HCI* 7(3). 315-346

Heritage, J (1984) *Garfinkel and Ethnomethodology*, Cambridge: Polity Press

Kraut, R E, Miller, M D, and Siegel, J (1996) 'Collaboration in performance of physical tasks: Effects on outcomes and communication', *Proceedings of the CSCW 1996 Conference on Computer Supported Cooperative Work*, 57-66, New York: ACM

Kraut, R., Fussell, S. & Siegal, J. (2003) Visual Information as a Conversational Resource in Collaborative Physical Tasks. *HCI Special Issue : Talking about things*. 18 (1&2). 13-49

Kuzuoka, H, Kosuge, T and Tanaka, K (1994) 'GestureCam: A video communication system for sympathetic remote collaboration, *Proceedings of the CSCW 1994 Conference on Computer Supported Cooperative Work*, 35-43, New York: ACM

Kuzuoka, H, Oyama, S, Yamazaki, K, Suzuki, K and Mitsuishi, M (2000) 'GestureMan: A mobile robot that embodies a remote instructor's actions', *Proceedings of the CSCW 2000 Conference on Computer Supported Cooperative Work*, 155-162, New York: ACM

Luff, P, et al (2003) Fractured Ecologies: Creating Environments for Collaboration. *HCI Special Issue : Talking about things*. 18 (1&2) 51-84

Martin, D. & Rouncefield, M. (2003) Making the Organization Come Alive: Talking Through and About the Technology in Remote Banking. *HCI Special Issue : Talking about things*. 18 (1&2). 111-148

Ou, J., Fussell, S., Chen, X., Setlock, L., & Yang, J. (2003) Gestural communication over video stream : supporting multimodal interaction for remote collaborative physical tasks. *ICMI'03*. ACM. 242-249

Sacks, H (1992) *Lectures on Conversation, Volumes I and II*, Edited by G Jefferson, Malden, MA: Blackwell

Schegloff, E A (1968) 'Sequencing in conversational opening', *American Anthropologist*, 70, 1075-95

Schegloff, E A (1972) 'Notes on conversational practice: formulating place', in Sudnow, D (ed), *Studies in Social Interaction*, New York: Free Press

Whalen, Jack (1995) 'Expert systems versus systems for experts: computer-aided dispatch as a support system in real-world environments'. In Thomas, P., (ed) *The Social and Interactional Dimensions of Human-Computer Interface*s. Cambridge University Press. 161-83

Whalen, J.; Vinkhuyzen, E. (2001) Expert systems in (inter)action: diagnosing document machine problems over the telephone. In Luff, P., Hindmarsh, J., & Heath, C., (eds) *Workplace studies: recovering work practice and information system design*. Cambridge University Press. 92-140.

Whalen, J.; Whalen, M.; Henderson, K. (2002) Improvisational choreography in teleservice work. *British Journal of Sociology*. June; 53 (2). 239-258.