From Tele *Presence* to Human *Absence* – the pragmatic construction of the human in communications systems research

R.Harper
Microsoft Research Cambridge
JJ Thomson Ave
Cambridge
+44 1223 479824
r.harper@microsoft.com

ABSTRACT

This paper reflects on the views of the human that were oriented to in two distinct research labs and which have been used to populate an inventive landscape over the past twenty years. It suggests that there are commonalities to the views in question, making them essentially the same. Both emphasise body movement at the expense of expression and both, one could reasonably claim, derive from a conceptual dualism as regards human nature associated with Descartes and then adopted by the computer scientist, Alan Turing. The paper will argue that, whatever conceptual dualists in philosophy or computer science might want to claim or emphasise, the use of this view by the researchers in question was not because it offered an adequate ontology but because it was a pragmatically useful way of looking at the world that enabled and helped drive inventiveness. The paper will report on how this was applied in the domain of communication technologies, particularly telepresence type systems. It will remark on the benefits and limitations of this view for the inventiveness in question and how this view led to many technological innovations that have not been widely adopted and to an indifference to innovation in textually mediated communication, amongst other things. The paper will remark on the value this view might have for future research.

Categories and Subject Descriptors

H.5.m [Information Systems]: Information Interfaces and Presentation –*miscellaneous*.

General Terms

Design, Human Factors.

Keywords

Sociology, ethnography, model(s) of the user; communications systems, media spaces, shared whiteboards, instant messaging, telepresence; computer science; philosophy, metaphysics.

1. INTRODUCTION

In 1904, Edouard Estaunie coined the term *telecommunication* by merging the Latin word *communico*, to impart or share, with the Greek word *tele*, for distance. He had in mind a word for any technology that used electronic signals to exchange information. We still use the term in pretty much the same way one hundred years later, though in a more encompassing fashion: the historian of telecommunication, A. A Huurdeman, for example, uses it as a label for any ‘technology of information transport’ – one assumes that papermail would fit into this category, as well as motorcycle couriers [14]. Somehow, though, we don’t think of telecommunication as quite the word for our communications obsessed times. One might say that the word *computer* and more especially the words *the internet* do instead; these words have become rough and ready synonyms for the all too human desire to be in touch. What I am thinking of here is not the evolution of words but simply the use of everyday phrases like “Are you on the internet?” and, “I am so much better at keeping in touch with a PC”. What these phrases convey is that some ‘thing’, a PC, a mobile, the internet and so on – whatever it might be – is like an umbilical cord linking each and every one of us; an umbilical cord between ourselves and our friends, ourselves and the world at large.

In this paper I am not interested in asking whether this is good for you or not (a question I do address elsewhere [11]) but I am interested in trying to define the kinds of vision of the human, the user, the one desiring to *be in touch*, that underscores the inventiveness that has produced this landscape. I want to answer this by offering what some might call an auto-ethnography, though this sounds rather pretentious for what is better described as a recounting and description of the thinking about users in the research world I have operated within over the past twenty years or so. I want to use my own contexts, if you like, to illustrate what has been the vision of the human communicator that has produced technologies that let us be in touch almost all the time.

When I say my contexts, I mean, in particular, a set of places that have been concerned with inventing various technologies, many of which address problems of communication. At first glance, the world I am thinking of might not seem the one that has been driving the communications landscape around us. One of these settings is Microsoft Research, for example. Now, though this place will obviously have had a great influence on what one sees on most PCs, it might not be thought that it has also helped create the communication world that we now live in. The other setting I want to report on is Xerox PARC’s

© The Author 2009.

Published by the British Computer Society

European lab, EuroPARC. Here too one might imagine that the research carried out was about office life and not about *being in touch*. But I want to say that in practice these places were (and in the case of the former continue to be) very much in the business of inventing the current landscape (along with other places of course). This is because these landscapes were and are being constructed, it seems to me, by what one might call a passion for computer-mediated communication and, underneath that, a vision of the user that has been used to motivate, justify and orient that passion. Of course this passion has been bound up with the emergence of the Web and especially the underscoring set of protocols that allow data traffic to move around the world at enormous speed and in huge volumes. But it is the human in this passion, the ‘interacting human’ that I want to focus on here.

This vision, I will suggest, has been an essentially dualistic one, with a distinction being made between external behaviours and internal intention, between the body and the mind. This dualism has a long conceptual history, being associated with the enlightenment philosopher Descartes, amongst others. I will suggest that, whatever its intellectual history outside the labs in question, the form that this dualism took in the workplaces I describe was *Turing theoretic*. When I say this, the view I will be putting forward may not be the one the reader expects. I will not be arguing about a vision of the human that has to do with intelligence, for example. After all Turing is famous for his so-called test of the same. This entailed asking a machine (a Turing machine) the same question as was put to a human and, if the machine did as well as the human, or its answers could not be distinguished from those of the human, then the machine would be deemed as having passed the test and ought to be treated as ‘equal’ to (or indistinguishable from) the human. Though one might think that with these arguments Turing wanted to encourage a concern with only one of this pair, namely the mind (where, in his view, ‘computation’ or something like it would be undertaken), what I want to propose is that in the research worlds in which I have lived, the opposite emphasis has been adopted. The emphasis has been *on the body*, not with the mind. In the view I want to recount, it is the *body* that can be seen as delivering the ‘intelligent act’ for example, just as it is the body that undertakes the communications act.

Furthermore, I will want to argue that this view has fostered what one might say is almost a whimsical indifference to important properties of words and their conceptual frameworks, of human action in the general, and human communication in particular. When I say whimsical I mean a casualness and lightness of thought. One of the reasons why I want to say the view in question is Turing theoretic (albeit seemingly inverted as I say), is because Turing himself had this lightness of thought, a lightness of thought that followed on from his basic Cartesian dualism. He had a casualness when it came to the word intelligence, for example; he also had a similarly casual indifference to the distinctions between machine-like behavior and the non-machine action when he sought proofs of intelligence. Likewise, my colleagues and I tend to show remarkably little concern for the mind as a linguistic conceptual category, for example, and even less with its correlate, ‘mindfulness’, for another. What gives credit to my use of the label Turing theoretic is then not that the theory in question is robust, delicate or comprehensive. It is rather that it is an orientation, and one aspect of that orientation is a remarkable shallowness as regards humanness and all that might mean from

the ability to express through words to the ability to be, occasionally, machine-like.

I need to make it clear that in claiming this I am not saying that members of my world have been or are unaware of the debates about, say, the phenomenology of action, a view that emphasizes the mind as embodied in action. As it happens, there has been a flourishing of interest in this topic over the same period of time that I am interested in (within HCI deriving in large part from Winograd’s work on the Coordinator [25] and in AI in the work of Agre [1], amongst many others). Similarly, many of the people in the landscapes I describe are aware of the possibility that the dualism itself might be egregious – a conceptual muddle, as Wittgenstein would have it [26]. The point I will be making is that these concerns didn’t matter to the individuals or the institutions in question, to my colleagues and I, since we had another focus, another concern, when we addressed our work. We weren’t thinking about philosophy, and though we might have adopted a kind of dualism, we did so because we were first and foremost interested in invention, on being inventive.

Before I start exploring what this view might entail or how it might be deployed I want to pre-warn the reader about what I will want to offer as insights from my investigations. For example, one lesson I might want to emphasise is that the view in question treats human behaviour as much like the behaviour of any animal, and not especially or uniquely human – it’s all about body movement and nothing else. But it seems to me somewhat commonplace to note that such a reductive view is empirically wrong, and can indeed lead one to misunderstand what is *human* about communication between humans. The real question, or rather a more interesting one, is what this view offers the inventive imagination and what are the limits of this view *on that count*. Answers to this aren’t to be in terms of what is a better philosophical ontology – they are to be in terms of what is invented. There might be various ways that this can be determined, of course, but one of them might be found by looking at the fate of the systems that the Turing theoretic view led to. As I shall show, one of the things that this vision led to was an emphasis on video (or the visual) in communication and less concern with text. Another consequence of this vision was the paradoxes and ironies that manifested themselves as ‘users’ of the technologies in question (and by users I will be including both myself and colleagues as well as ‘users’ outside the labs) were confronted with things invented for the body (so to speak) and yet used those same things to do things that are neither for the body or for the mind. The users appropriated the technologies in question for a world in which that separation didn’t (and still doesn’t) apply – the real word if you like, not the world as seen from the inventors’ prism. As it happens, I shall note that, in some cases, these appropriations have been quite consequential for the communications technology landscape around us, leading to the widespread adoption of applications that were byproducts of the inventive ambitions rather than key targets of those same ambitions. In other cases those appropriations have disappeared and been lost to the historical record, and instead of being used to create new visions that might supplement or colour the visions used to invent, are simply neglected – hardly written about, rarely reflected upon, not confronted for what they are: efforts by people to make technologies do things for them, even if those things are not what the inventors hoped or expected.

2. Interacting individuals

I want to start my inquiries with a vignette of things I was involved with some twenty years ago. I want to describe two or three systems that we were toying with then, say something about our expectations as regards what they let us do and then say something about their fate. I then turn to describe how I still see colleagues, today, playing with variants of some of these concepts (by concepts I mean an invented thing, a device). I won't want to say that here is proof that my colleagues and I don't seem to get anywhere but I will want to say that what motivated us then is still pretty much what motivates us now.

So, to begin: in the early Nineties, I was fortunate enough to work at Xerox EuroPARC, here in Cambridge. EuroPARC was a curious place to work, in lots of ways, and one was the predilection of many of its staff for video-mediated communication or VMC. One consequence of this was that the EuroPARC building – the entire building – was wired for sight and sound. More or less each and every office had a coaxial cable linking it to every other (there was a lot of cabling!) and these conveyed high quality audio-video connections between offices. Thus what was called a 'media space' between any two or more offices could be created. Part of the fun of working at EuroPARC was the shock and glee that the media space provoked. Some journalists were appalled, thinking it Big Brotherish; others enamoured at how the media space was pointing towards the prospect of being able, in the future, to bring people separated by great distances, closer together within their office worlds.

Needless to say, many projects were undertaken investigating the utility of the system – or rather set of systems – that comprised the computational infrastructure of EuroPARC's media space. In my mind, the most enjoyable aspect of this was called Kronica [16], though it was the media space *qua* media space that got the most attention [3. For a general history see 13]. Now, looking back, a couple of projects are of note.

The first had to do with shared editors or what were most often called

shared whiteboards. These allowed people in different offices to type text into a document at the same time as someone else, elsewhere, in another office. These shared editors worked as follows: one of us would type text into a window on a screen (the 'editor window' of the application); our colleagues in another room would see this being typed into their view on the editor window and would then, if they so desired, type on the next line of that window. This would be visible to the person who created the first line. Typically, one could not tell who wrote what unless one of us typed their name at the end or – and this was more common – one could tell simply by the turn-taking. I wrote, she wrote, I replied; that was the gist of it.

Of course, our actual use was a bit more complex than this. Also, the reader may get suspicious that they are familiar with contemporary applications that sound very similar: they will be right. But before I come to that let me explain some more about what we wanted with these editors. These applications were being investigated for their use in what was called knowledge work. Whatever knowledge might be – it's a rather vague term at the best of times – we were interested especially in supporting such work when the participants were separated by distance. The thought was that new organisational structures were increasing the likelihood that such work would be undertaken by individuals spread around the world. Organisations were

going global; knowledge work was being done in New York, London and New Delhi; technologies to support distributed knowledge work were required. Though our office was only small while the distances of the global organisation enormous, it was thought that the distance between our individual offices could be treated, analytically, as proxy for bigger ones in the real world. The shared editors or whiteboards were meant to exemplify this philosophy.

But, as it happens, nothing much came of these editors, whatever the merits of investigating knowledge work. By nothing much I mean we did not manage to implement them in any effective way in 'real work', that is to say in our work or in other's. When I say manage, I am perhaps being disingenuous. Though these tools clearly were central to what one might have said was our research remit, we did not find them hugely interesting. They seemed altogether too prosaic to be worth investigating (though I should say some visiting researchers did find interesting aspects to them which they published in the groupware and CSCW literature: one particular editor was researched extensively not only in our lab [See 15]). Our managers remained vaguely hopeful that we might evolve these shared editors into being novel work tools, bringing a Xerox sheen to what was a fairly crude technology. But we did not. Other technologies, especially the audio visual aspects of the media space, enticed us away.

Before I say why, let me note that nevertheless we used these shared editors for ourselves. But we did not use them for knowledge work – to help us write reports or share profound ideas. We used them to play. We used them because they became devices for laughter and mischief; because using them could be fun. The topics were mostly jocular, asides about the burden of deadlines, occasionally about plans for after-work activities. At least one romance flourished at EuroPARC with this form of exchange. I am mentioning this now because what we were playing with then was the genesis – albeit in a convoluted and indirect way – of what we now know as instant messaging – as the reader may have recognised. What our managers thought would be tools for professionals and their knowledge practices, indeed what my research colleagues and I dismissively assumed would be work-orientated applications, ended up being crafted to support, some twenty years later, such things as teenagers keeping in touch with each other when they have been 'imprisoned' in their bedrooms by parents hopeful that once there, they might then do some homework. Of course, I am making fun of instant messaging and the wiles of teenagers [See for example [9,10], but my point is a serious one. Tools for serious things like work have ended up being used as tools for play in social life.

Why we believed that these applications were to be thought of only in terms of work seems a little perplexing now. After all, we did not use them to support our own work even though we were professionals doing knowledge work of a kind. As I say, we found fun and utility in them through their playfulness, in supporting not professional engagement but our social selves – we were doing then pretty much what teenagers do now. One might pursue this line of argument and suggest that we were indulging in what the French sociologist Maffesoli suggests, metaphorically, in his *The Time of the Tribes* [17], is a contemporary form of tribalism or more specifically, a *Dionysian tribalism*. He has in mind the desire of our 'modern selves' to celebrate the temporary, daily, bodily togetherness of work tribes, whichever tribe that might be – research lab tribes

included [7,8]. I think there is certainly something in this metaphor, but it is not appropriate to develop it here.

2.1 Who is the user that is invented for?

Going back to my concern, with the 'kind of human' underscoring our inventiveness, the fate of this technology can lead one to ask a number of questions. One might ask, for instance, whether there was something peculiar in the relationship between those who did the inventing and those for whom the inventing was said to be done. Was this an instance of inventors justifying their ideas on the grounds that someone, if not themselves, would use their inventions when it turned out these some 'ones', these 'ordinary people in the real world', were something of a construct; even artificial, one might say? This leads one to think of the term Otherness used to label the concern of anthropologists: one might say that one of the ways that we work in my world is by making users 'other' in various ways and here, this otherness was simply other than ourselves.

Certainly I would be the first to admit that, over the years, I have indeed heard my colleagues justify some idea(s) with the assertion that 'real people' will use them, even if the colleagues in question (the inventors of the thing) don't say who these people might be. I have made the same assertions myself on occasion. It is probably true to say, also, that my colleagues, just like me, can drift into laziness; tiredness even and so don't really have in mind a very particular 'real world user'. And besides, phrases like 'ordinary people', 'real people' and 'people in the real world' do sometimes get used as a kind of bland catch-all to justify technologies that my colleagues and I haven't really figured out the purposes of ourselves. We think we have devised something that is of use, but for what and for whom is beyond our comprehension. Ordinary people have lots of faces and lots of desires and so we use the label, 'ordinary user' (or its synonyms) as an expression of hope, if nothing else, a hope that someone, somewhere, for some reason or other, will indeed like our contraption(s).

Going back now to the historical record at EuroPARC, it is certainly true to say that in some instances we didn't have a clue who might have used our inventions, but this was not very common. By and large my colleagues and I were pretty thorough in our work and most often did indeed have a view of the 'user' in mind. This view, earnestly held, seriously thought about, was not, I suggest, a view of the user that one might hold outside of the work place I am describing. I want to suggest that this view, not always articulated in-depth it has to be said, was nevertheless the view that the majority of my colleagues assumed they could 'work on'. More importantly, just as they held it then, I want to suggest that my colleagues in another lab hold it now.

These examples of shared editors have only hinted at what this view is. Before I can say something about how that view plays out today, let me now uncover some more aspects of that view – making it clearer – by recounting another technological experience afforded by the EuroPARC media space. Our absorption and fascination for this will tell us a great deal more, I think, about the kind of user we had in mind. As we get a better sense of that, we will also begin to see, I hope, what this vision consisted of and how it motivates us in many ways even to this day.

2.2 Ways of 'looking' in my world

When I had first arrived at EuroPARC, the media space system was fairly simple. Each office had a camera – a high quality camera – mounted on a wall or tripod, or attached to some

shelves. It was through this that others, using the media space, could see you. There were speakers, too, for sound connection and these would be fitted somewhere in the room where space could be found. There would be a monitor as well, and it was this that would let one see the 'other' office when a connection was made. All these were separate from the workstations we were using – though we used the workstation to initiate a 'connection'. One consequence of this arrangement was that in most cases all these devices were separated from one another. Hence if one were to look into the camera then one might well be looking away from the monitor on which one could see one's colleague. The same could hold true for one's colleagues. So, when one made a connection to another, one would look at them on one's own monitor but they would appear to be looking elsewhere – not at you. They too were looking at their monitor wondering why we were looking away. In both cases, the camera would be some way from the monitor – perhaps not far – but far enough for this 'incongruity of perspective' to occur (as my colleagues expressed it).

Some of us found this vexing or at least a cause of mirth. 'Look at me when I am talking!' they would shout; 'I am!' we would reply. Some of my colleagues, though, were convinced that this was problem that needed fixing. They devised inventions of sorts that solved the problem. Doing so was not difficult – somehow the cameras needed to look at the subjects as if the cameras were looking from the same point of view as the monitor. Both the looking and the glancing needed to be the same. At that time, in the early Eighties, the only way of achieving this 'fix' was to use a system of mirrors, like a periscope, assembled and boxed up around the monitors such that the line of sight of the camera appeared to be hidden within the monitor. The system – or rather this improvement to the system – was called *video tunnels*.

Once built, various experiments were undertaken to investigate whether 'users' (us) felt more at ease with the system. One particular concern was to see whether users felt better able to understand what others in the system were looking at. With this system, one knew that when someone looked away they weren't looking at the picture of oneself, for example, and therefore, might have demonstrable purposes with some object, a document say, that one might orient to oneself. There might be a 'mutual referent' as it was grandly put. These findings were published. But at the same time, once the experiments were completed, the system became somewhat neglected. The periscope type boxing was somewhat large and unsightly; besides none of us seemed to care for the difference video tunnels made. Within weeks neglect turned to vexation, and the set up was dismantled and the media space returned to its former state.

Be that as it may, from my own view, it certainly seemed easy to use a system when the eyes of the fellow participant seemed to be looking at oneself, but at the same time much of my own enjoyment with the media space had nothing to do with the ability to look at another's eyes as they gazed into one's own. Whether one was using video tunnels or not, it was fun merely to peer into another room, to see what was happening there. Sometimes my colleagues and I would use the media space to simply call out, between rooms separated by several floors, and see if someone was around. This was especially so between reception and the admin offices, on the bottom and top floors respectively. This play did indeed seem a kind of Dionysian tribalism, as Maffesoli would have it; all the more so as I look back now and wonder at how much we played: I was younger

then of course, many of my colleagues were too: EuroPARC was a young lab.

I am not wanting to comment on that playful tribalism, though. As I say, I am interested in the view of the user that motivated this enhancement to the media space. I want to get even clearer on that view by recounting how, now, many years later, my colleagues, here, at another research lab, albeit in the same city, are endeavouring to invent something very similar. But this time, they are devising something without the ugly boxing, the heavy camera or the co-axial cables. Whereas then we had to make do with very large cameras, now video cameras can be very small indeed. Moreover, then the best we could do was stream the video data over the network, now we can process it and interrogate it. This makes a considerable difference to what can be assembled, or as I am wanting to call it, invented.

The new system I have in mind combines two very small video cameras attached to the top of the monitors at either end of the link (in two offices, say). As before, the cameras do the looking, while the monitor's display what is to be seen. Thus far this new system and the one in the media space are the same, albeit in size and numbers of cameras. But here the similarities end. In the new system, developed under the auspices of a program of research called *i2i*, data from the cameras are processed with so-called intelligent algorithms so that the system can 'recognise' and follow the movements of those looking at the monitor. In simple terms the cameras produce stereoscopic data that the computer can analyse so as to identify where the user's head is. There are various ways of doing this, from tracking movement of the head itself (as a three dimensional outline) through to tracking of specific points on the head, such as the eyes, tip of the nose and mouth. Irrespective of the tools (or 'keys' if you prefer) the system uses, the result of this arrangement is that, once the tracking is processed, the computer can adjust the image shown on the remote monitor to ensure that when the user is gazing at their own monitor, it looks as if they are gazing through the cameras at the remote viewer. The actual discrepancy between the location of the cameras and the screen is in effect dissolved. The end result of this system is quite delightful, it has to be said, albeit slightly odd. Odd because one knows that the remote person's eye's are not really gazing at you; they are looking at a screen displaying a virtual you. But delightful too since it affords a kind of playful paradox: it's nice to be looked at, but after a while one tries to see how far one has to go before one cannot be looked at – i.e. to go beyond the processing of the system so that it cannot 'correct'.

What was the goal behind these two systems? In one way I have said it already and it seems obvious, indeed, perfectly reasonable. When one is having a video conference (or videocall), it would be nicer if the person you were conferencing with appeared to be looking at you, rather than somewhere else. Nicer seems a weak word here, but it does convey the sense of greater politeness and social grace that the systems afford. The point I am wanting to make though has to do with the relationship between my colleagues ideas of what to design for, ideas related to what the human actor or user of the system was needing – and how to deliver, through design, this sense of grace. It is certainly true to say that people, when together or when using a video conferencing system, will somehow manage to co-ordinate their gazes and orient their behaviours in such a fashion that they produce for each other a sense that they are 'in touch', 'of a mind', and 'doing something collaboratively'. It is certainly true also that there is

a great deal of empirical literature on how a sense of 'jointness' is achieved in social action, conferencing especially. This is especially so in the ethnomethodological canon. But somehow our research wasn't really about how this synchrony of looking would make for courtesy. Systems of social grace and propriety were not the motive behind the design of the systems in question. Rather, what was behind the systems was what one might call an idea about the *body mechanics* of interacting humans. I am thinking here of designs which sought to allow lookings and glancing to 'fit together'.

2.3 Fitting touches

I need to be careful here because there is an extensive literature reporting on the moral dimensions of looking and glancing – researchers like Egon Bittner have explored how *not* responding to a look can be seen as social rebuff, for example, just like *not* answering a hello. His paper *Must We Say What We Mean* points out that meaning is conveyed even when we don't speak or even look [2]. This is because of how our actions embody our intentions and are thus visible for others to see. David Sudnow does the same [23, esp. pp259-280]. But this argument (that meaning is incarnate in our actions) is not what I am wanting to pursue now. What I am wanting to suggest is that what we devised in my workplaces was a much simpler vision of human interaction that somehow emphasised just the body mechanics of the interaction and somehow lost sight of this moral dimension.

I now want to refer to another system my colleagues have been developing, more recently, also a conference system of sorts, and show that this has a similar concept of 'body mechanics' as the key of human communication behind it. This view is persistent, I want to say. If the systems I have mentioned thus far have concerned themselves with the 'how' of looking faces, the system I now want to describe concerns itself with the touching and moving of hands.

The system I have in mind, *C-slate* [4], uses a stereo camera and some clever object recognition software to notice when hands are placed on a digital document, when they point, for example, or even when they appear to erase a word. The cleverness, the inventiveness, is making such movements recognisable to the system. It is made possible by developing, first, the ability of the machine to discern shapes, in particular the shape of a hand and a finger in this case, from other shapes and forms that the cameras might see. Second, this is combined with the capacity to determine movements of these things when the movement in question indicates something. Hence, a finger's movement can be seen as a gesture, for example, not a mere shadow moving over a document. The result is a system where users can get to grips with their digital documents. They can use their hands, moving on top of a document, to mark out, edit, remove, paste, as they so desire, words, images, stuff if you like, in the document. The movements of their hands, the gestures of their fingers, and so on, are converted by the machine into what one might think of as equivalents to mouse movement and pointer clicks. Such a technique is not devised so as to replace the mouse or the cursor, as it is a method for interaction over documents that are being interacted with by people over distance. One person points and shakes his finger over a document; a person somewhere else can see this interaction and can have the consequence of this interaction made manifest in the document they are working on. This is not merely, then, a picture of a hand on a document, it is a hand *interacting* with a document that itself is digital; and a document which is itself shared or distributed.

In other words, we have in mind the idea that, whatever it is that people want to do, they can be aided in that task by systems that replicate what one might call the *physical geography* of the interactions in question. In the first case, in the media space examples, we imagined that, if we were designing a communication experience for people who want to communicate to those far away, then one way of supporting that would be to invent a system that would let the people in question glance and look and gaze at each other when they are remote just as they might when they were together. What we needed to design for was what one might call the *interactional geography of lookings and glancings*. In the other example of the more recent system, we were also concerned with remote communication, but in this case, interaction and communication about documents. Here a system was invented that meant that how someone uses their hands when dealing with a document could be conveyed, replicated to someone remotely. Here it is *the geography of touch* that mattered.

The way I have put it, as geographies of touch and lookings, makes our vision of the human and of human needs in communication technologies seem obscure, even arcane. This is not my purpose. My mannered description is designed to make strange what is in fact commonplace in my world, the world of inventing computer-mediated communications systems. Making it strange can help us think about it, I hope. What I am suggesting is that what motivates a great deal of the inventiveness I see around me, as illustrated with the enhancements to the systems mentioned above, is a concern with bodies, with the human body, or more particularly, with the bodily mechanics of human interaction with computers and hence, also communication through and with computers to other people. I am saying that inventiveness in my world is about devising ways of allowing the mechanics of communicating individuals to fit together. In this vision, it's not what is said in communication that matters, nor why something might be said, nor what may be the consequences of the saying of something. It is in the *how* of the saying, the manifest behaviour entailed when people communicate. When using the word 'said' I am obviously evoking words, but as the examples above should attest, in this vision, it is not just words that matter. What it is that is exchanged in a communication might be broad indeed – glancings, pointings or gestures, just as much as the sharing of the written or the spoken word. In all cases, I am saying that I and my colleagues sought in the past and seek now ways of letting the body mechanics of those doings be replicated. Replicated so that they are conveyed over distance, for example; replicated so that the assembly of those mechanics, the sounds, the glancings, the pointings, may be conveyed as an 'assembly'. And I am thinking too, of how the interaction of multiple, communicating bodies needs to be replicated – glancings need to be synchronised; listenings need to be two-way; distractions ought to be shared if they are to be oriented by another body's lookings and glancings.

2.4 A vision of the human user

Now, a number of questions and queries naturally follow on from this. One obvious question is why this view? Why not another? How does this view affect or drive invention? Will there come a time when bodily communicative acts can be entirely replicated, for example? This seems unlikely, but what are the difficulties here? Besides, the examples I have used seem rather confined, essentially to do with knowledge work and conferencing of various kinds – so what about the daily assault of messaging and callings and blog postings we see

around us today? These don't seem like body mechanics; so where have these come from – what inventive landscape produced them? Aside from these matters, what about motivating inventiveness: is satisfying this 'ideal' (of bodily fittings, irrespective of whether it is right or wrong) motive enough to keep my colleagues and I at the mill stone? Could other visions of the human perform a similar role: namely the role of keeping researchers passionate enough to keep on inventing 'this' and inventing 'that'?

One can't answer all of these questions, though some are certainly more important to our concerns. Let me deal with the question of where this view of the human comes from; this can provide a basis for commenting on other questions like motivation. It would appear that my colleagues and I have a vision of the human, the user if you will, and in particular a vision of the actions of the human, which one might say is not really very common – ordinarily we don't think of fitting our bodies together when we communicate (or rather sometimes we do but not often). Though, as I said earlier, it is sometimes the case that my colleagues coin the phrase 'ordinary user', it seems to me it is not a very ordinary vision of people they have in mind. Following on from the formulation that Peters uses in his history of the concept of communication [19], I think there is a metaphysics here, a set of ideas about who and what the human communicator is. In this case these derive from various sources in philosophy, and, in turn, via their interpretation by computer scientists. As I said at the beginning, I think one might say Alan Turing, one of the founding fathers of this discipline, might be accountable here. It seems to me that his influence relates to his encouragement of a view that splits the human into either body or mind, into a dualism, and then tends to encourage a whimsical approach to what one of the part in that dualism might be, namely the mind. Turing showed little interest in and hence sensitivity towards what the term mind or its correlates (such as intelligence) might mean, nor what mindfulness in action might entail. One consequence of this is that those who have acted in his wake, in workplaces where the metaphysics of computer science are often commonplace, such as it was at EuroPARC and now at Microsoft Research, is that there is even less attention given to mindfulness. More attention is given to the apparent vehicle of the mind and intention, namely the body.

As it happens, there might be other ways of reading Turing and the detailed claims he made in his various papers—there is no mention of bodies in his work for example – but it is I think a fair comment on the *consequences* of his view in the places and processes I am thinking of – in institutions dedicated to the art of invention, concerned with technologies for human communication that have helped fabricate the age of *being in touch*.

Nevertheless, it is perhaps worth pausing a moment to see what he did say. In his various papers on 'intelligence' and the 'computer machine' written in the Thirties and Forties, Turing claimed that one may understand intelligence by treating it as to be solely measured in external performance, in behaviour (see for example his 1936 paper, '*On computable numbers, with an application to the Entscheidungsproblem*', loosely meaning the 'halting problem') [24]. Whether a machine or a person performs that behaviour was irrelevant, he proposed. This had methodological advantages for him, offering, as he saw it, a route around the problem of moral judgement. A person could not be relied upon to objectively describe a person's action as intelligent or otherwise for fear of upsetting the person being

judged. Thus, in Turing's perspective, only behaviour must count and the source of that *behaviour* must be hidden from view. The fact that in so doing Turing reduced the concept of intelligence from being rich, complex and subtle, used in diverse ways to achieve different sorts of understandings, purposes and descriptive values in ordinary language and life, into something that is little more than a label for a binary opposition, where an action is either intelligent or not, seems not to have worried him. I don't want to pause and comment on that at any length, here, though (but see [20]). Of more importance is that this move has another consequence, as follows. Those who adopt Turing's view assume that what goes on inside the machine itself is not only invisible, it is somehow *tricky* and *best avoided*. The 'moral implications of judgement' or, rather, the moral aspects of a person's thoughts are such that one ought to steer clear of them if they are to get anywhere.

Now, as it happens, this hasn't meant that followers of Turing's metaphysics haven't subsequently looked at what goes on 'inside the head' and attempted to do a kind of science. But when they have, they have done so from the presumption that one starts with bodily behaviour, with body-ness if you will, and have to construct, with external evidence, ideas about what goes on inside the head. For some of those who have taken this Turing theoretic view seriously, the implication is the possibility that if a machine, a *computer machine*, can equal a human through the use of a programme (i.e. some lines of code as in a Turing test), then perhaps it is the case that the *human mind* is computer-like too. In their opinion, one should research to discover what that code that makes-the-mind-work might be. In computer science this view is held by the likes of Marr, who claimed that the way a computer machine can recognise a visual object (via various forms of Bayesian prediction, for example) must be the way a human mind processes and hence 'sees' [15]. Outside of computer science, this idea has been adopted too, with much of contemporary cognitive science being a product of the possibility that one can divide the human into the body and mind and assume, also, that the mind is a processing machine, just like a computer (but see [19,20]). One could go on. Turing has been enormously influential in all sorts of ways. But here I am not wanting to offer a general history: I could explore how Turing's view can be seen to have its roots in earlier ideas – before Cartesian dualism, to when the mechanism of clocks provided inspiration for how to grasp what the human was [See 6, esp. pp233- 257]. But that is not my concern. I am wanting to recount the inventiveness I have seen around me in terms of a vision of the human that has underscored it. I am suggesting that a concern with body mechanics has motivated us and I claim that this is so not because I think we have adopted that concern out of intellectual investigation into philosophical ideas about 'humanness'. If, as the sociologists might have it, we have *configured* our view of the user, we have done so first and foremost for pragmatic reasons [5]. When I say this I mean that we have sought visions of human endeavour and the human actor performing those endeavours that have prompted ideas that could lead to invention. The body emphasising view I have described has helped us to achieve this. This is what a Turing theoretic view has come to be in the workplaces I have described: a pragmatic philosophy that highlights 'problems' and 'needs' of communicating humans, bodies seeking lookings and touching, glancing and pointings, all of which let my colleagues and I conjure technological solutions to what we thought we could see as problems of human contact, of being in touch.

In saying this it might appear that I am suggesting we have adopted a poor or enfeebled vision of the human. This may be the case from some points of view. But this is not what I am arguing. The view in question ought to be judged by what it let us do. I want to suggest it led us to what we thought were complex designs, not simple designs. In these respects, this view worked for us¹. That is to say, the examples show that the credit one ought to give to the model of the human we used is in terms of what it led us to invent.

So, as a case in point, one of its values was that we thought it would lead us toward more complex technologies than other views on humanness might have done. Thus we turned away from early versions of instant messaging because we thought them technologically prosaic and mundane in terms of the richness of the (*human*) communication in question. We looked instead to audio visual systems. As we did so we thought that we were addressing more *complex and profound* issues of human communication. Our view led us to focus on complex bodily matters and thereby not on matters that might be measured in terms of human value or intention. As remarked, we used the shared whiteboards for some fairly complex social actions – play, romance – but viewed those behaviours as less interesting to invent for than media space interaction because they used only *simple technologies* and from our pragmatic standpoint only *simple behaviours*. In our judgment, they used only one mode of bodily action – namely the keyboard. Our other inventions, our enhancements to the media space, seemed more appealing since they used multiple modes of bodily action as well as more complex technologies. What appealed to us was precisely more sensual dimensions in communication. This appealed because more dimensions of this kind would demand more innovative technological solutions, we thought. In short, though our vision of the human may seem enfeebled, in fact the contrary was the case: we thought and often still do think that the vision leads us to see greater complexity: greater sensual range in bodily behaviour, greater richness in the sense of collaborative symmetry in distributed work tasks, for another.

2.5 Squeezing humanness into movement

One has to be careful when one judges such motivations, then. From some views, say a social scientific one, the metaphysics in question might look impoverished, but from another, in this case from the point of view of those using the metaphysics, the opposite may seem true. A pragmatic vision used for invention is not to be compared with the view of an empirical science. They are doing different things.

The measures that one ought to apply are ones that are, if you like, internal to the goals in question. Thus, to judge the vision of the user I have described would lead one to look at the inventions it produced; this seems to be one such relevant 'internal' criteria. As should be clear, this vision did drive us to invent and this inventive spirit that it enabled did indeed produce a lot of 'IP' as the saying has it in corporate life (IP standing for Intellectual Property). This IP focused on the manifest bodily behaviours of expressing people and had to do with technologies that enabled the geographies of movement to be telemediated.

Our vision led us to focus on these concerns by leading us away from focusing on other things. We did not worry about things that could not easily be rendered in terms of body mechanics, for instance. Nor did it foster much interest in the activity that is most distant from bodiliness: namely the written word. Though written words are created by the hand, the values of

words are in their meanings, in the intangible ether of language, not in the mechanics of the body. Text-based applications might have been built with the metaphysics I have described in mind but they were built as vehicles to get to higher order things, to better audio-visual systems in the media space, not as solutions in their own right. Our indifference to what became instant messaging is a case in point.

This vision also made and continues to make us neglectful of our selves, of our own activities, leading to paradoxical relationships between what we invent and what we use. As we saw, when the video tunnels experiments were completed, we packed up the enhanced system and went back to the old audio visual set up. We did so because we didn't imagine we would find it useful for ourselves, in our own activity. This was because what it offered, a richer set of bodily fittings, did not actually equal an easier mode of expression. Indeed, quite the contrary. When one used the video tunnels, for example, the question of bodily fittings, of mutual glancings and lookings, was only part of consequences or implications of the actions in question. When I shared a video connection with others, those others weren't primarily interested in the particular objects I looked at; they wanted to know *why I wanted to glance at them at all*. To make a video connection was an enormously powerful act, an assertion of a kind of intimacy, a closeness that was not something that each and every one of my colleagues thought 'appropriate'. By appropriate here I am alluding to the moral implications of such acts, not their physiological ones. Yet such concerns, though 'real' were beyond the scope of our inventive landscape. We did not worry about these concerns. Besides, if we did bother to look at them closer (over coffee, say, or when we reflected at the end of the day) we saw how tricky the concerns in question would be to unpack and use as a basis for design. Sure, we recognised that there were issues, concerns, topics, properties if you like of human action that were relevant here, but our metaphysics encouraged us to avoid them.

If we had problems with fitting the mind's landscape to the audio visual system at EuroPARC, to the bodily landscape the system enabled, then we have even more problems fitting the latest systems we have devised to allow interaction over distance. At least at Xerox we played with the infrastructure, but with our new interaction-rich, eye focused conference systems such as i2i or our new document centric C-slate, we don't even play. At Microsoft, once experiments are completed they are packed up and put away. None of my colleagues have any desire to use these systems for themselves. Again, they use the phrase 'ordinary people' to suggest that others, somewhere else, might be willing to do so. But why they don't themselves want to use such systems is something that they view as uninteresting and to be avoided.

3. Conclusion

My purpose in this paper has been to claim that we have a notion that what we are doing and where we are going does have some sense or purpose; direction even. My colleagues and I are afflicted by a conviction that somehow, to someone, our colleagues, our managers (perhaps), and above all, to the user, 'out there' in the 'ordinary world', what we are about does indeed have reason. This hope, and our personal convictions about it, has led us to invent not just anything, I have suggested. I have explained that it has led us to explore technologies of a particular kind, based around a particular set of topographies of what is understood as 'being human'. I have argued that this

has been one that emphasises the doings of the body, what I have called the *geographies of interaction*.

The examples I provided were intended to indicate the direction of the inventive imagination behind research. The direction leads us along a path where, crudely speaking, 'more' is viewed as better; though what is meant by more is rather particular. One can think of the shared whiteboards I mentioned earlier on to illustrate what I am thinking of. These applications allowed persons to write collaboratively, over distance. But we saw these as impoverished, and as being something that could be improved by offering applications that allowed something more, but here we did not think of more written words. We thought of sight – and hence our imaginations turned to the media space. More meant a move from *text* to *sight*. And if we were entranced by sight then so, today, we are entranced with touch. We saw this with C-slate.

These examples are meant to show how this credo leads my colleagues and I to invent applications, devices and technologies, that are essentially designed to allow the body, the communicating body, to do more. Though we might have been fumbling in our efforts to do this, we have been persistent nevertheless. Even today we are working on ways of using tactile input and output modes to extend the range of sensualities that systems afford; we are seeking ways of offering 3D renderings of objects in video mediated work which are clearly, in our view, an improvement over 2D ones – they must be, we believe, it makes more sense given our body emphasizing credo.

Oftentimes, in HCI, debates about how to configure the user appeal to ideas about what is empirically right or wrong. Tacit in these debates is the assumption that more evidence is better and that the model of the user is 'scientific'. But what I have wanted to show is that HCI is also a pragmatic endeavor, and that one should not muddle scientific verisimilitude with the basis of inventive imagination. In saying this I have not also claimed that what fosters inventive imagination may not feed off empirical evidence nor be subject to change. It is with this in mind that I now want to ask the following questions. As we look back – as I look back – now, over twenty years of inventiveness, what do I think ought to be the next step, the future of the inventive imagination? One can see that the credo my colleagues and I have used has been productive, to be sure, but one can see too that it has forced us to neglect some things. So, should my colleagues and I alter it? Should we abandon it altogether and come up with a new *credo*?

In *Speaking into the Air* [19] Peters suggests that our recent fixation of bodies in telepresence type systems has had the paradoxical effect of making the body become more important in human expression than hitherto. Whereas text-based systems (by which he means the written letter that had its heyday in the Victorian era) tended to emphasise the mind, he says that telepresence systems have tended to emphasise the body and in so doing have brought an erotic charge to the communications act. It is the sensuality of the body that users have come to seek when they communicate in new modalities, and it is these self-same sensualities that inventors try to convey in their technologies, he believes. It seems to me that we might learn from Peters. If we see our concern with bodyness as one that is expressly about the erotic, then we might be better able to see what our endeavors have ended up offering as well as seeking, and we might also see in greater clarity what our endeavors have missed and ignored.

For one thing, when we use the word *erotic*, one comes to see that the idea that work might be done better when people's bodies are brought together virtually seems a little odd. If it was the case that we were originally interested in media space and related technologies because they might help change the functioning of spatially distributed knowledge work then, as we look back now, we can see that that hope might have been quixotic. Certainly I do remember thinking that with our inventions new forms of work might emerge in which spatial dimensions might become less salient, less of a problem if you will. But that media space type tools have not become ubiquitous since that time suggests to me that the spatial properties have remained resistant. The maxim that lies at the heart of Adam Smith's *Wealth of Nations* (published in 1776) still seems to apply, more than two hundred years after it was written: people work better when they are separated from one another. Of course this doesn't apply to all work, as much of the research in CSCW (in command and control rooms in particular) attests, but in the general this would appear to be the case. A second and closely related issue has to do with how our research credo emphasised the synchronous in communication. Body fittings were to be done in real time, in our view. This has led us to explore the possibility that people need to work on the same task at the same time. But, again, this seems to contradict most work practices which, though undertaken concurrently, typically entail doing different jobs. Adam Smith can be evoked again here: most work is achieved successfully through breaking it up into a complex matrix of interdependent tasks sequenced through space and time. This doesn't mean that people don't do things together at particular spatial and temporal locales; after all most people have many meetings to attend and these are certainly that. But even meetings may not be about doing the same things together in the same place and time; they might be about monitoring what others are doing so that each and every attendee at a meeting can go back to do their distinct task confident that in the end all of their work will tie up together. Meetings have many other functions, too, and many are not to do with work itself, but with ceremony, with corporate identity and much else beside. Work places aren't only about doing the work, but about feeling part of gang. One could go on. The bottom line is that though we might have had brave intentions when we started on our course of invention, time might be suggesting that those intentions were too bold. Perhaps today we might reconsider how we configure our user because of these lessons.

With this concern in mind, one might use the emphasis on the *erotic* to help us better understand what has not been achieved more clearly too. We might learn from that how we might recharge our inventive imaginations. For example, we can begin to account for those uses of telepresence systems that have not been reported in the literature. As I say, media space connections at EuroPARC were very much experienced in terms of the illicitly *erotic* through the sense of intimacy they provided. It was hence that it had to be treated very carefully. Yet this was never written about. A large part of what media spaces afforded then was nothing to do with work and work activities, but to do with the Dionysian that I mentioned – and all the constraints and delicacies that that has in a workplace. As I say, when a media space connection was made, the issue that arose was not how much easier it was to do work, but why someone would want such a contact if only work was at issue. When seen in this light, media spaces aren't about work, but about touch, not reason, about passion, not a distribution of labour.

But one might learn from this. For if such technologies cease to be about work and are about the geographies of interacting bodies, about the *erotic*, then the places in which those technologies might have their place, and the activities that they might support, might be different too. If the word *erotic* helps us grasp what might be enabled by the technologies supporting body interaction then we might broaden our conception to include the delicious wonder that bodies can afford. Hence one can think of grandparents wanting a media space so as to get a sense of the bodily presence of a remote grandchild for example; one can imagine the appeal of a media space connection to enhance family events too – special dinners, birthdays and so on – when bodily presence gives them especial power.

Of course the *erotic* is only one part of family life. But if the use of the word *erotic* helps define a space for inventiveness perhaps the starkness of what the word implies will also highlight other possible avenues of research, other credos that will foster inventiveness for the mundane, the sterile and the prosaic – all the things that aren't *erotic* if you like. One of the problems of the Turing theoretic credo I have been portraying is that it offers a kind of autistic vision of human endeavour, one that cannot countenance things beyond its scope – like mindfulness and textuality. Perhaps the credo of the *erotic* might be better precisely because in emphasizing one thing it also highlights its opposite: what the body affords is touch as well as aloofness, the pleasure of being together and the delight of a written word sent from afar; absence as well as presence.

4. ACKNOWLEDGMENTS

To my various colleagues over the years, as well as to anonymous but helpful reviewers

5. REFERENCES

- [1] Agre, P.E. (1997) Towards a Critical Technical Practice: Lessons Learned in Trying to Reform AI, in Bowker, G. Star, S.L. & Turner, W. (Eds) *Social Science, Technical Systems and Cooperative Work*, Lawrence Erlbaum, Mahwah, NJ, pp131-157.
- [2] Bittner, E. (1977), 'Must We Say What We Mean?' in Oswald, P. (Ed) *Communication and Social Interaction*, Grub & Stratton, New York.
- [3] Buxton, B. Moran T. (1990) 'EuroPARC's integrated interactive intermedia facility (IIF): Early experiences', *Proceedings of IFIP WG8.4 Conference on Multi-User Interfaces and Applications*, Heraklion Crete.
- [4] Izadi, S. Agarwal, A. Criminisi, A. Winn, J. Blake, A. & Fitzgibbon, A. 'C-Slate: Exploring Remote Collaboration on Horizontal Multi-touch Surfaces', in *Proc. IEEE Tabletop, 2007*
- [5] Grint, K. & Woolgar, S. (1997). 'Configuring the User: Inventing New Technologies,' in *The Machine at Work: Technology, Work, and Organization*, Grint, K. & Woolgar, S. (Eds), Polity Press, Cambridge, pp 65–94
- [6] Hacker, P.M.S. (2007) *Human Nature: The Categorical Framework*, Blackwell, Oxford.
- [7] Harper, R. (1992), 'Looking at Ourselves: An Examination of the Social Organisation of Two Research Laboratories,' In *Proceedings of CSCW '92*, 31st Oct.–Nov. 4 Nov. (Toronto), New York: ACM, pp. 330-337.

- [8] Harper, R. (1996), 'Why People Do and Don't Wear Active Badges: A Case Study', Harper, R.H.R. In *CSCW: An International Journal*, Vol. 4, pp.297-318.
- [9] Harper, R. (2005a), 'To Teenage Life to Victorian Morals and Back: technological change and teenage life', in Glotz, P. Bertchi, S. Locke, C (Eds), *Thumb Culture: the Meaning of Mobile Phones for Society*, Transcript Verlag, Bielefeld, pp101-113.
- [10] Harper, R. (2005b), 'The Moral Order of Text: Explorations in the social performance of SMS' in *Mobile Communication – Perspectives and Current Research Fields*, Hoflich, J. & Gebhart, J. (eds), Peter Lang GmbH – Europäischer Verlag der Wissenschaften, Berlin, pp199-222.
- [11] Harper, R. (Forthcoming), *Texture: Human expression in the age of communications overload*, MIT Press.
- [12] Harper, R. & Hamill, L. (2005), 'Kids will be Kids: the Role of Mobiles in Teenage Life' in Hamill, L. and Lasen, A. (eds) *Mobile World Past, Present and Future*, Springer-Verlag UK, pp61-73.
- [13] Harrison, S. (Ed) (Forthcoming), *Media Spaces*, Springer Verlag, Goldalming.
- [14] Huurdeman, A.A. (2003), *The Worldwide History of Telecommunications*, Wiley. New Jersey.
- [15] Hymes, C. & Olson, J. 'Unblocking brainstorming through the use of a simple group editor', in *Proceedings of CSCW 1992* (pp. 99-106). ACM Press, ACM Press.
- [16] Lovstrand, L. (1991) 'Being Selectively Aware with the Kronika System', (Bannon, L. Robinson, M. & Schmidt, K. Eds) *Proceedings of E-CSCW 91*, Kluwer, Dordrecht, pp249-264.
- [17] Maffesoli, M. (1996), *The Time of the Tribes: The Decline of Individualism in Mass Society*, Sage, London.
- [18] Marr, D (1983) *Vision: A computational Investigation into the Human Representation and Processing of Visual Information*, Freeman & Co, New York.
- [19] Peters, J. D. (1999) *Speaking into the Air: A history of the idea of communication*, University of Chicago, Chicago.
- [20] Piccinini, G. (2003), 'Alan Turing and the Mathematical Objection', in *Minds and Machines*, 13: 23-48.
- [21] Shanker, S. (1998), *Wittgenstein's remarks on the foundations of AI*, Routledge, Abingdon.
- [22] Sheutz, M Ed, (2002) *Computationalism: New Directions*, MIT Press, Boston, Mass.
- [23] Sudnow, D. (1972), *Studies in Social Interaction*, The Free Press, New York.
- [24] Turing, A. (1936) 'On computable numbers, with an application to the Entscheidungsproblem', *Proceedings of the London Mathematical Society*, Vol 42 (1936-7).
- [25] Winograd, T. (1987) 'A Language/Action Perspective on the Design of Cooperative Work,' *Human-Computer Interaction* 3:1 (1987-88), 3-30.
- [26] Wittgenstein, L. (1953) *Philosophical Investigations*, (trans A.N. Anscombe), Blackwell, Oxford.

ⁱ One might say that Turing's view of the human in the 1936 paper also allowed him to do *his* work, namely to define some parameters of computer algorithms. That was the focus of his thinking, not some ontology of the human. As it happens, in later years Turing did drift in to that concern but that is another matter.