Telepresence as a Design Philosophy

An Interview with William Buxton

William BUXTON

Interviewer: MORIYAMA Kazumichi

It's Not Technology, It's Interactive Design

MORIYAMA Kazumichi: Could you begin by telling us how you define the term "telepresence"?

William BUXTON: First of all, many people think about technology in the sense of interaction between people and machine. But actually I think the most important thing is technology-mediated communication from human to human. So telepresence for me is one of the most interesting research areas. And it's a good idea to be precise about what I mean by "telepresence" because many people have many different meanings for telepresence.

In some places, telepresence is "tele-robotics" or "tele-manipulation." While in my work—my experience is in the area of social interaction, i.e., human-to-human interaction, over a distance—telepresence is having a strong sense of presence. But presence of what? It's presence of the person, presence of the activity, presence of the artifacts. To have complete communication, it is important that all of these things are there, both physical and virtual. In this conversation we have already shared many things, but that is very difficult to achieve in, for example, video conferencing.

The second part of what's important for me—and this is where many problems remain—telepresence is not just going across distance in geography, but also in time. So that a person can participate in a meeting even if he cannot be actually there. But the question is, can we go over time, and not just distance, with the same technology? Today you have video conferencing for going across distance and you have maybe VCR for going across time. But they are never integrated and never considered the same problem. Yet in many ways they're the same problem.

Another thing that's very important is that with different professions, such as computer animation compared to automobile design, for example, the artifacts are very different. So you need a very different kind of technology for the conversation, because what must be communicated or scanned or printed is so different.

Although we have made some progress in telepresence between two locations, it's very difficult to have multiple people in communication among more than two locations. It's a big problem. But the real problem is that they are being studied as a problem of technology. It's not a technology problem.

If my arm gets chopped off, I go to the doctor and get an artificial arm: that is a physical prosthesis. Technology for me is a social prosthesis. If I lose my arm, you have to know what I do with my arm before you make an artificial arm. You need to know what I do—I like to work at my computer, I like to draw. When you know my activities, then you can make a design that fits my needs. Your interests may differ, so a design for you might be very different. I have to understand your activities to design technologies to give you capabilities.

Support via telepresence is also prosthetic: we need to understand the nature of meetings and the nature of interaction. It's not a technology question, but a question of sociology. If we look at most telecommunications companies or people investigating these topics, maybe 80 to 100 percent of the researchers are engineers. In my opinion, that should be at least 50 percent social scientists.

I can draw a figure here. Let's pretend that each of these are people in an organization. In our office here on this floor everybody can tell just by looking who's there. Every time I talk to this person or talk to that person, either by email or telephone, I draw a line. Now, this looks like a telecommunications network but it's actually a social network. If there is much communication, I could make it thick. But I can also make a geographical representation of the social interactions. In fact, there's a mathematics called "graph theory," which studies the properties of this kind of diagram. Every telecommunications engineer knows how.

But if I change from telephones to just people, we forget that we can use the same technique to analyze the flow of information. So here's how I can design a technology to get people in the organization to have the desired structure. While I need the technology to make the design, I also need other knowledge. I must have some notion of "health"—the notion of a healthy sociology.

In many cases, when people make these technologies, even the language they use is the wrong language. They use the word "communication"

because they think it's the transfer of information that's important. But the important thing is trust, right? So I call these technologies "trustification," not "communication." It's not in any dictionary; it's my buzz word.

Three Skill Levels

WB: We can look at how to do these things. I need to know what kind of social organization I want for the organization communicating. I need to understand things about them—how often, what artifacts and where they are in distance and where they are in time—so there's a sense of community. Communication is not the goal; communication is the means to community.

So what kind of community do we want? That's the framework for how I think about this intellectually. Then I approach such things through what I call "ecological design." Which maybe is not a common term. Ecological design, in my definition, is design that takes into account the physical, social, cognitive and psychological ecology of the people who will use the design.

The first notion when I look at design is the skills people worked very hard to get that gives them their identity and their pride. For example, I worked very hard to learn how to ski. It's a skill I have. When I buy skis, I expect the technology to respect the skill that I spent 30 years to develop. I studied music for many years: I was a professional musician. So when I get a saxophone, I expect the designer of the saxophone to respect all of those days I spent ten hours practicing, all of the techniques I developed. There must be respect; the tools must fit the skill.

Everything I've mentioned so far fits my motor skill, but actually there are three levels of skill in every human: you have motor-sensory skill (how your hands, your ears work); you have cognitive skill (how you think); and you have social skill. Every one of those skills is different: the motor skills of an athlete are different from the motor skills of an artist, which are different again from the motor skills of a musician. The cognitive skills of an accountant are different from the cognitive skills of a philosopher. If you design a technology to fit a cognitive mental model, then you can have a very successful product. For example, Lotus 1-2-3, or before that, VisiCalc: the reason spreadsheet software was successful is the technology respected skill and how people thought about accounting. It was designed to the right model, and it was immediately successful. Now we see the same thing: different societies, say, Japan compared to Canada. Even an animation company in Tokyo like Dream Pictures, which is making feature films using my company's software, has a very different social structure from Honda, which is designing automobiles also with our software. We must understand even those cultural differences.

When I look at any technology, I think about it as a mirror. The quality of the technology is in how well it reflects my skills. In fact, there are three mirrors: quality in reflecting my cognitive skill, quality in reflecting my motor-sensory skill, and quality in reflecting my social skill and behavior. Ecological design takes into account this notion of how design reflects the skill, but also the skill in the physical context: not just the person, but the person's location. The goal then becomes to provide the right function at the right time for the right person in the right form, in the right design. Since these factors will change, the technology for one person will be different than for another.

The Technological Pursuit of Divergence

MK: So we need to design special technologies to meet diversified scenarios--to be used where and how and by whom....

WB: Yes, and that's where we have problems today. Today every computer looks the same. If I look at the technology—video conferencing and the telephone—all of these technologies are exactly the same if you're an accountant in Japan or a computer scientist in Germany or a medical doctor in Canada or a computer animation person in France. In fact, none of those technologies have changed in concept for more than 15 years. If I look at a modern computer with Windows 95 or NT, it looks exactly the same as Xerox Star that came out in 1982. If I look at video conferencing today, it is no different from video conferencing 20 years ago. If

I look at telephone today, there's been no progress. The only progress is in cost and distribution and maybe speed. But conceptually, the idea behind the design philosophy has not changed.

We must be smarter than that. There's a famous old Broadway song called "Is That All There Is?" When I look at technology today, I ask, "Is that all there is? Did the people who invented these technologies get it correct the first time? Will it get no better?" My answer, of course, is I hope so, because it's terrible right now. So what I have to say is, first, what's needed?—the social side. Second, what's the design approach? Now we can move on to Part Three—what are some examples of how these things can happen?

Let's talk about telepresence. If we want to make a new technology to support collaboration, maybe it's a good idea to look at an old technology to support collaboration and see what we can learn. I don't think the problems are new. The technology is new, but the problems are not—which is actually a good thing, because if it's all new to us then we can say, "I'm not educated. I don't have any bases to make decisions right or wrong." If I'm designing technology and everything is new, then I'm lost. But designing social structures and organization structures—that's not new. All of a sudden I can smile, because I've got knowledge, I can understand it. So let that be the basis for designing. There have already been technologies to support existing organizations.

Now, here's a very good exercise. Take a piece of paper and keep a record of every conversation you have during the day and make a note of how long it lasted. Did you plan to have the conversation? Was it scheduled? How many people were there and where did it happen? It's very interesting to notice that some take place in your office, some take place in a conference room, some take place at lunchtime in a restaurant, some take place in a corridor because you bump into people, some take place outside—let's say you take a walk outside because you don't like your colleagues. Some take place on email, some take place on the telephone. Then if you note what you talk about, you start to understand that we choose the location depending on why we're meeting. We also notice that we don't always have meetings in the same place. There's a rich variety of meetings and structures designed to accommodate that rich variety.

Which brings me back to technology-mediated interaction and telepresence: what do we do if we have one video conferencing room and all meetings—no matter if they're scheduled or unscheduled, serious financial meetings of the company directors or brainstorming between junior engineers—take place in that one room? When in history did we ever believe that one location could support every kind of meeting for every kind of person in an organization? All technology design more or less follows that approach—certainly in video conferencing—with a central location, a central room. It's the same as saying that every meeting in this office must take place in the same room. It's crazy. Everyone would quit the company the next day.

It doesn't matter how nice this room is, how much everything cost—this is a video conferencing room!—or how expensive the video conferencing equipment is. The important thing is to distribute the technology so that there are as many different meeting places for electronic meetings as there are for face-to-face meetings. They should be different not only in location but in the kinds of meetings that they support. We have a wealth of different meeting environments, so they correspond.

So we come back to Part One: what kinds of meetings do you have? Where do you have them? What are the parameters? Why do you have them in that location? Why do you have them in the boardroom instead of in an office or in the corridor? Once we start to understand what differentiates that location from this location, then we can start to think about how to design the technology that seamlessly supports that kind of meeting.

The Disappearance of the Computer

WB: That brings us to Part Four. Everyone is saying that we are now in a period of "convergence": telephone, computers, everything is coming together into one thing. Convergence is the story—everyone is telling us—of our salvation. In fact, it's exactly the opposite. Convergence is the worst concept we can possibly imagine.

Now I have to be careful about what I mean by this. If you

Ire a plumber, convergence is a good thing. Or, let's come back to buildings. In this building there is a central network that brings water to this building, to the taps and so on; it also takes the dirty water away. Instead of the Internet, let's call that the Waternet. The Waternet is a centralized, standard, converged network of water. All the water, whether it's for washing or for swimming or for plants, comes from there. But let's take a look at how many different water appliances there are in the building, how many different kinds of sinks or toilets. You know if you're in a men's washroom or a women's washroom because they are different. You know a sink for washing dishes is different from a sink for washing hands, which is different from a sink for washing clothing. With sprinklers, you can tell if it's for a shower for you, for the grass in your garden or for a fire in a building. There are many different appliances, all of them specialized; there's divergence there. That is very interesting: it's very much in contrast to how computers today exist. If water appliances were designed the same way as our computer appliances, if we would drink, swim, pee and wash in the same sink, again, that's crazy.

The most important thing today is to make the computer disappear. Consequently, we must recognize that convergence is only for the plumber; it's only for the Waternet area. The real value only comes when you have specialized information appliances that are distributed in such a way that you know this room is for laundry and this room is for showers and this room is for dishes. Or in a business, this room's for large meetings and this room's for small meetings, this room's for engineering meetings or for financial meetings or for casual meetings or for lunchtime, and so on. All of a sudden we realize that divergence leads us to what has been called "ubiquitous computing."

The problem with ubiquitous computing is the word "computer." Everyone thinks they know what a computer is. If I ask you to draw a computer, everyone will draw a computer terminal, not a computer. The computer is down in the basement. This is what I want to get rid of. I want to have as many different kinds of terminals as water appliances everywhere in the house, to develop that kind of network.

Specialized Technologies

MK: Not convergent, but divergent, decentralized, with only the input-output devices close at hand--in different forms according to various different uses. This thinking comes as a natural extension of ecological design. How, then, does your idea of "ubiquitous computing" express itself in terminal device design?

WB: The basic idea is to make computers reflect the way people work. So if I make something for an artist, I'll make something that looks like a desk, a drawing surface. If I'm a teacher in a classroom, my computer should look like a blackboard. Or, in your house you have a refrigerator and there are probably magnets on it for holding pieces of paper. On my refrigerator I have pictures of my children, my horse, a calendar with my family, messages from my wife, pictures that my children drew, all these things. What's interesting is that what's on my refrigerator is actually a homepage. I only need to put something on the front of the fridge and it goes into the computer. So imagine your refrigerator where you take a picture, you take a pen and write a message, everyone living in the house has a little index. You can also check email, your voice mail, your faxes, all right there on the fridge. Now imagine I'm in traffic and I transmit a message "I'll be late" to my refrigerator, my wife will see the message when she comes in, across the front of the refrigerator because that's where we leave messages.

For telepresence and collaborative work, and collaborative living now, I want my refrigerator on the Internet before I want my computer on the Internet. For me the refrigerator is more important because that is the real information appliance in the household. It's also a good example of ecological design. Already the calendar is there, already it's in my family—the social structure uses that location. Why would I change that to put a computer in the kitchen? It makes no sense.

MK: So you're saying we should have specialized technologies to suit particular places and the activities in those places?

WB: Yes, technology knows what it's supposed to do. Now imagine you go to the grocery store and buy food for

the week. You walk out of the grocery store with both hands full. The door of the grocery store knows it's not a bank. In a bank the door would be very secure, but the door in the grocery store knows the sociology of the grocery store: it knows you probably have both hands full and no free hand to open the door.

Then here's a very interesting situation. [takes out a digital camera] This is a computer; this is not a camera. This computer has light-in–instead of keyboard-in–and chemistry-out–as opposed to an LCD panel. The important thing with this computer is that it knows its function. It knows it's not a word processor, it's not a telephone, it's not a radio, it's not a spreadsheet, it's not an animation program; it is for taking pictures. Because of that it has a knowledge of light, exposure, focus and all this knowledge inside. And the power of the microprocessor inside is probably greater than an Apple II computer. And because of how it's designed it's very easy to use. So whether we're talking about computer design or video conferencing, I like to talk about this camera as an example. With, say, a Nikon F or a Hasselblad, it's like MS-DOS or UNIX–you can take any picture you can imagine with that camera, but the probability of the picture coming out well is very low unless you are an expert user. Now when I look at photography, there's only two basic questions that matter: what and when. They translate to two actions: point and click. That's all. I can change the focus and do other things, but they're second-level decisions that I'm not forced to make. So what used to be very complex is now very simple and the best thing is the probability that the photograph is a good one is very high, because this device knows its identity; it knows what it's for. It doesn't think it's a traditional camera. It's a hard drive where it can write from a CPU. And this CPU happens to have a different input channel.

This relates to "ubiquitous computing" and a change from just talking about ubiquitous computing to what I call "ubiquitous media." All of what's important in ubiquitous computing is also true in video conferencing. They all should work together. Now I have my telecommunications technology and my computational technology working together in specialized locations for specialized purposes. As with the camera, I put intelligence in for the intended task. The more I know about the purpose of the device, the more intelligence I can put in it and the less it has to do–it's narrow, but deep. There are so many tasks, each one different. That's the real benefit of divergence.

Human-Human and Human-Machine–Foreground and Background

MK: Then, telecommunications media should also diversify to serve specialized functions?

WB: Yes. Another important aspect: if we look at technology, there's actual (technology-mediated) human-human and human-machine interaction. The telephone, video conferencing and email are human-human interaction in the foreground, up-front, while electronic whiteboards and GUI (Graphical User Interface) are human-machine interaction in the background. Almost all research so far has been in the area of foreground; research in the background has been missing.

An example of human-machine, background interaction would be the door in the supermarket. The foreground activity here is "take the groceries to the car." The secondary activity is "walk" and the third is "the door opens for you." In this, human interaction with the machine is the same as in washrooms when you put out your hand without having to touch, because the tap knows your hand is there. You know what's funny? The toilet is smarter than my Thinkpad, because this computer doesn't know I'm here. It can't react. It doesn't know that I'm not alone. It doesn't know I'm showing it to people. But with a little camera, it would be easy for it to know that. The sink in the bathroom knows that, why does this computer know nothing about its context, where it works? It's stupid. It has millions of switches and the toilet with only one switch still knows I'm there. It would be funny if it wasn't so sad.

What is human-human communication in the background? An interesting question, isn't it? Think about it: in Japanese offices mostly everybody works at a table like this, in the open–everybody is in one big room. When I'making drawings, you can see what I'm doing. While working on my own job, I know what other

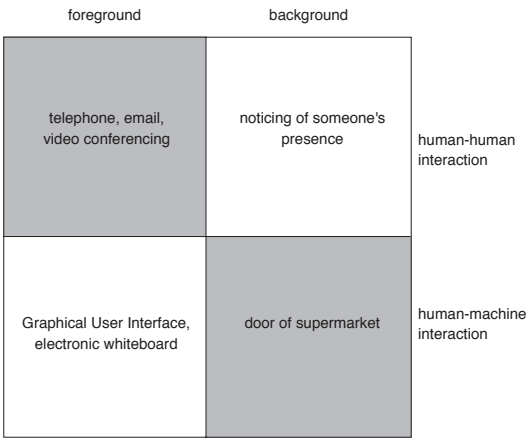
people are doing; I know who's there. When I have a problem, if I see you, I can grab you and say, "I need you for a minute. Will you help me?" So, like the camera—humans are like cameras—the foreground action I take is point and click. The background action is focus, expose for the light, and all these other things. Likewise, my foreground is to do my drawing, but in the background I know who else might be important, and when you come by, I can quickly leave the foreground, look up and ask my question, then quickly make the transition back again.

MK: I see. So in that instance would you see "presence" as a sense of backgrounded awareness of actually being somewhere?

WB: Well, noticing maybe. All the research so far has focused on what comes after you approach someone. Whereas most meetings are in fact accidental, not planned. Yet there's no research on how you recognize an opportunity, say, walking around on the same floor in a building, and how that makes for a meeting. The telephone is very good in the foreground, but I never accidentally meet you on the telephone. You must either call me or I call you.

This is where we worked at my laboratory at the University of Toronto and at Xerox PARC. When I'm in my office in Toronto, I keep small images on my screen of the offices to be connected. It just takes one small photograph per minute to see or hear you in the background while I'm working. And when I find the right moment—bang—we connect.

In telecommunications, recognizing opportunities necessarily comes first. And the information that lets you know this must be visible at all times. Second, you must have the means to translate that opportunity into action, to make the connection. Only then can you begin to have the kind of social interaction we're having right now.



William BUXTON
Chief scientist, Alias / Wavefront and Silicon Graphics. Associate Professor at Toronto University

MORIYAMA Kazumichi
Born in 1970. Science writer and producer.
URL: http://www.moriyama.com/