

# 1

## Human skills in interface design<sup>1</sup>

**William A. S. Buxton**

*University of Toronto & Xerox Palo Alto Research Center*

### 1 Introduction

In an earlier work (Buxton, 1986), I speculated on what conclusions a future anthropologist would draw about our physical make-up, based on the tools (namely computers) used by our society. The objective was to point out that these tools reflect a very distorted view of our physiology and the motor/sensory skills. For example, the near absence of pressure sensors reflects a failure to exploit a fundamental and well-developed capability of the hand. The impoverished use of sound reflects a waste of our ability to use audio to make sense out of our environment.

The paper dealt primarily with the domain of the visible and tangible. Nevertheless, things have changed very little in the intervening years. Furthermore, it can well be argued that things are even more distorted if we look at how the technology reflects less visible human traits such as cognition, or social interactions.

In what follows, we use a technology-as-mirror metaphor. One intent is to provide some *human-centred* criteria for evaluating designs. Another is to help foster a mind-set that will lead to improved designs in the future.

---

<sup>1</sup> Citation: Buxton, W. (1994). Human skills in interface design. In L.W. MacDonald & J. Vince (Eds.). *Interacting with virtual environments*. New York: Wiley, 1-12. An earlier version of this paper appears in Buxton, W. (1991). The three mirrors of interaction: a holistic approach to user interfaces. *Proceedings of Friend21 '91 International Symposium on Next Generation Human Interface*, Tokyo, Japan, Nov. 25-27, 1991. This earlier version was reprinted in English and Japanese in: Buxton, W. (1992). The three mirrors of interaction: a holistic approach to user interfaces. *Industrial DESIGN 157*, Japan Industrial Designer's Association, 6 - 11.

## 2 Three mirrors

The thesis of this chapter is that we should consider technology in terms of the fidelity with which it reflects human capabilities on three levels:

- *physical*: how we are built and what motor/sensory skills we possess;
- *cognitive*: how we think, learn, solve problems and what cognitive skills we possess;
- *social*: how we relate to our social milieu, including group structure and dynamics, power, politics, and what social skills we possess.

Our metaphor is one of three separate mirrors, each reflecting one of these levels. In order to be judged acceptable, designs must provide an acceptable degree of fidelity in how they reflect each of these three aspects of human makeup and activity. The benefit is in how the model can provide a simple but valuable test that can be used during the design process. We now look at each of these mirrors in turn.

## 3 Mirror One: How we sense and control

### *Look and feel*

The notion of *Look and Feel* primarily encompasses aspects of the user interface that are reflected by our first mirror. The term has had a lot of recent attention, largely because of the efforts of various manufacturers to protect their own approach. Looking into our first mirror, however, what we see is something distorted and unworthy of protracted protection.

First, the very term reflects a lack of concern with a primary sense - sound. And if we look at the interfaces in question, then something like:

look      feel      sound

would be a more accurate representation of what is actually there. Look dominates, feel is impoverished and sound, while used, is almost a "throw-away." In short, the balance is out of all proportion with human makeup and capabilities.

One of the first priorities of the next generation of user interface, therefore, is to correct the current distortions in mirror one. How some emerging systems and technologies can lead the way in doing so forms the basis of the next few sections.

### *Multimedia*

*Multimedia* is one of the topics that inevitably arises when discussing emerging technologies. As likely or not, the discussion has two main components: "Multimedia is the future!" and, "What is multimedia?" The results are generally more than a little confused.

I think that a lot of the excitement about multimedia is well founded. However, by definition, "Multimedia" focuses on the *medium* or *technology* rather than the *application* or *user*. Therein lies a prime source of confusion.

If we take a user centered approach, we quickly see that it is not the media *per se* that are important. The media are transitory, one to be replaced by another as technology evolves. What makes these technologies different are the human sensory modalities and channels of communication that they employ. Rather than focusing on the media, therefore, more appropriate and focused terms might be:

- *multi-sensory*: design that utilizes multiple sensory modalities;
- *multi-channel*: design that utilizes multiple channels, of the same or different modalities;
- *multi-tasking*: design that recognizes that (as driving a car demonstrates) humans can perform more than one task at a time.

Multimedia is simply design that makes better and broader use of the human's capabilities to receive and transmit information. However, since the vast majority of the literature has been technocentric, confusion has resulted. Without an understanding of the properties of the various sensory modalities, the associated channels, and the way in which they work together, how can effective use of the technology's potential be motivated or achieved?

Let's look at some examples that provide counter-examples to the mainstream, and point to what can be achieved when one breaks away from the *status quo*.

### *Bidirectionality of the Senses*

An interesting attribute of most human-computer interaction is that it uses a different sensory modality in each direction of communication: primarily *visual* from computer to human, and *motor/tactile* from human to computer. This is almost taken for granted. But it is contrast to almost all human-human communication.

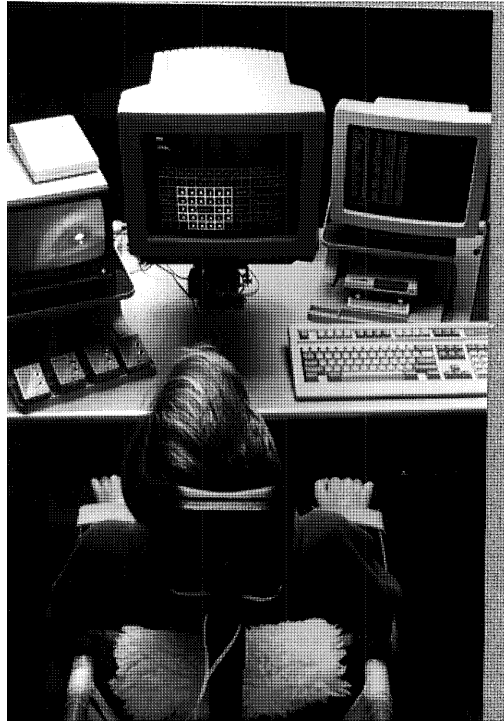
This is brought up because recognizing and balancing asymmetries is a good heuristic for uncovering new design alternatives. Let's take the visual system, for example. Everyday we use our eyes not just to see, but to direct the attention of others to what we are looking at. Here is a well practiced human skill which is a candidate for use in human-computer interaction. Here is an opportunity that systems such as that by LC Technologies Inc.<sup>2</sup>, shown in Fig. 1, have exploited.

Similarly, we think of the things that we touch and move as being input devices, only, and spend very little time thinking about their use as tactile *displays*. In one way, virtually

---

<sup>2</sup> LC Technologies, Inc., 4415 Glen Rose St., Fairfax, VA., USA 22032.

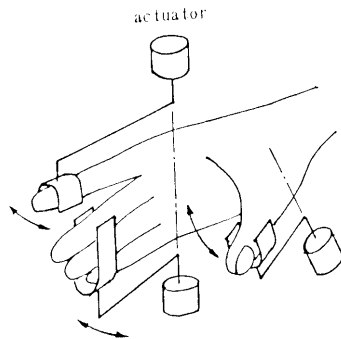
all input devices are also output in that we can feel their size, shape, action, and movement. But while important, this is generally passive display. Force-feedback devices are one approach to coming closer to taking advantage of our capabilities of touch. Examples of such devices are described by Brookes *et al.* (1990) and Iwata (1990), which is illustrated in Fig. 2. But much of the same effect can be achieved by smaller tactile displays, such as Braille displays designed for visually disabled users. Another example is the tactile display mounted in the button of the SimGraphics mouse, shown in Fig. 3.<sup>3</sup>



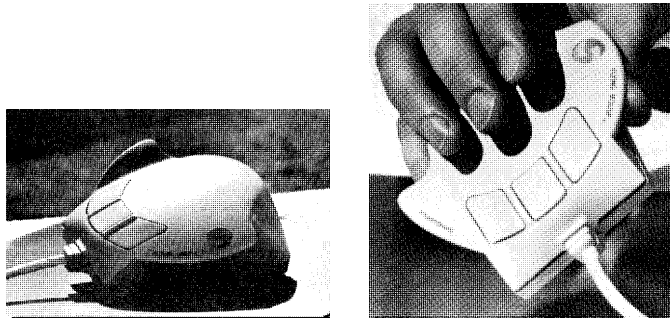
**Figure 1:** A Non-Intrusive Eye Tracker. A video camera mounted under the display tracks the position of the eye's pupil and translates the data into screen coordinates. Thus, the eyes can be used for pointing. (Photo: LC Technologies, Inc., Fairfax, VA).

---

<sup>3</sup> SimGraphics Engineering Corp., 1137 Huntington Dr, Suite A-1, South Pasadena, CA, USA 91030



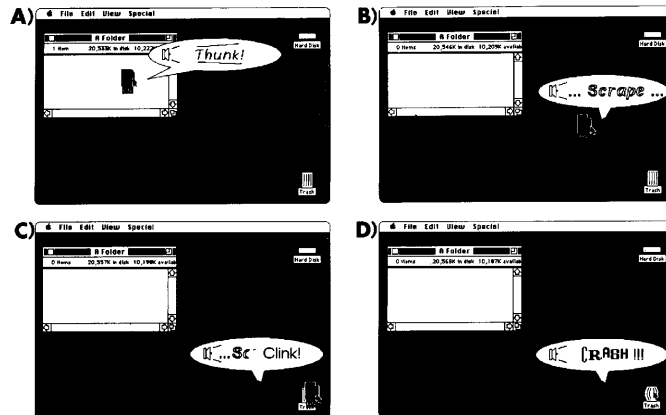
**Figure 2:** A force-feedback hand controller (Iwata, 1990). The device is a stationary exo-skeletal mechanism. The schematic shows the actuators that provide the feedback.



**Figure 3:** This 6D pointing device has an array of computer-controlled pins in the button that rise and fall under computer control in order to provide a degree of tactile feedback, such as to simulate the feeling of crossing an border, or coming into contact with an edge. (SimGraphics Engineering Corp.)

*The SonicFinder and beyond*

One of the most interesting pieces of software circulating in the research underground is something called the *SonicFinder* (Gaver, 1989). It was developed at Apple Computer's *Human Interface Group* by Bill Gaver. The *SonicFinder* is a prototype version of the Macintosh Finder that is based on the novel proposition that we can hear. This may seem a fairly obvious fact, until we look at the sonic vocabulary used by most computer systems.



**Figure 4:** The SonicFinder. The example illustrates the use of sound during four stages of dragging a file to the trashcan. (A) The user selects the object and hears the impact. (B) The user drags the object and hears it scrape along the desktop. (C) The user drags the object over the trashcan; the scraping sound stops and the trashcan makes its noise. (D) The user releases the object; a crash provides confirmation of the deletion. (from Gaver, 1989)

What the SonicFinder does is use sound in a way that very much reflects how it is used in the everyday world. You can "tap" on objects to determine their type (application, disk, file folder, etc.) and their size (small objects have high pitched sounds, large objects are low). When you drag an object, you hear a noisy "scraping" sound. When a dragged object "collides" with a container (such as a file folder, disk, or the trash can), you hear a distinct sound.

Now all of this may seem to suffer from terminal "cuteness" or frivolity. After all, aren't sounds for video games and cyber-wimps? But how many times have you missed the trashcan when deleting a file, or unintentionally dropped a file into a file folder when dragging it from one window to another? Frequently, I would guess, yet these are precisely the kinds of errors that disappear through such use of sound.

We have ears and are pretty good at using them. Machines that have the capability to exploit their potential are finally becoming more common. Starting with the Commodore *Amiga*, which comes with rich audio and text-to-speech, we are now seeing audio as an important ingredient in other platforms. The challenge now is to learn how to use audio effectively, not just for music or to provide some acoustic lollipop, but as a means to provide a sonic landscape that helps us navigate through complex information spaces.

### *Hands-on computing?*

Every day we turn pages with one hand while writing with the other. We steer our car with one hand while changing gears with the other. We hold a ruler or drafting machine with one hand and use a pencil in the other. All of these tasks employ the performance of everyday motor skills that have potential in human-computer interaction, but are largely ignored by computer systems.

Computers like the Macintosh were designed for Napoleon: unless you are typing, you can work all day with one hand tucked into your jacket. Fine if you are one-handed - a

waste if you are not. The image of the user reflected in the technology is lopsided to the extreme.

"Hands-on" computing is largely a myth. It would be better called "hand-on" or even "finger-on." If we are accurately to reflect human potential, we should be able to scroll through a document by manipulating a track ball with one hand, while pointing with a mouse using the other. We should be able to scale an object using a potentiometer in one hand, while dragging it into position with the other. Or, in a program like *MacDraw*, we should be able to move the drawing page under the window using a trackball in one hand, while still keeping our "pen" in the other.

To a certain degree, high end interactive computer graphics systems have used this type of interaction for a number of years. This is seen, for example, in systems that provide a potentiometer box which enables the non-dominant hand to perform transformations such as *rotate*, *scale* and *translate*. To date, however, such se f two-handed input has not penetrated the mainstream market. This is about to change.

*Can you deal with the pressure?*

Using two hands is not enough, however. Another property that we all have - but which is not reflected in the computer technologies that we use - is the ability of our hands to control and sense pressure. Gradually, designers are beginning to recognize this capability and design for it. One lead comes from electronic musical instrument keyboards. There - unlike your mouse - keys are not simple switches. Each key has what is known as *aftertouch*: the ability to sense continuously how hard the key is being pressed.

Soon (hopefully), *aftertouch* will be standard on mouse buttons, thereby providing natural control for line thickness, scrolling speed, or the speed of fast-forward or rewind on videos or CD-ROM. Already, a few manufacturers, such as *Wacom* and *Numonics* make styli for tablets that are pressure sensitive.

But no matter how well the look, feel and sound of a user interface is developed, it still may not fit how we think or work, and therefore fail. To gain a better understanding of this, and see how emerging technologies might help, we can now look into a different mirror.

## **4. Mirror Two: How we think and problem solve**

*The Myth of the Information Revolution*

There are would-be sages and futurists who will tell you that we are in the middle of an information revolution - the impact of which is matched only by that which followed the invention of the printing press or the industrial revolution. Unfortunately, this notion is false, and illustrates a serious lack of understanding of the nature of information.

By definition, information is that which *informs*, and which can serve as the basis for informed decision making. Rather than an information revolution, what we are currently experiencing is a data explosion. The combined advances in contemporary telecommunications and computational technologies have helped to spawn an era where true information is ever more difficult to find at all, much less in a timely manner.

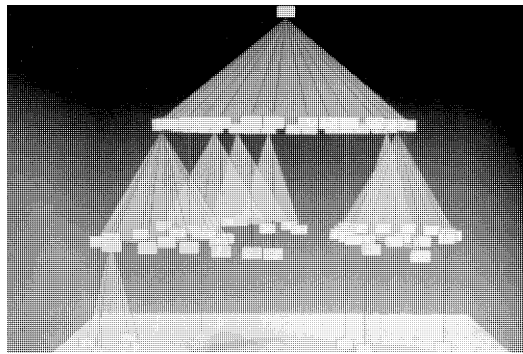
Information technologies that deserve the name are less computational engines than technologies to filter and refine data into a form where it does inform, and warrants the name *information* (Tufte, 1983, 1990). Just as we want systems to reflect how we hear, see

and touch, we want those same technologies to accurately reflect and support how we think, learn, problem solve, and make decisions.

One of the biggest successes of the personal computer world is the spreadsheet. One of the main reasons for this is that it "fits" the way that people think about certain problems. Rather than generate masses of new numbers, it helped users refine data into information by enabling them to explore and understand new relationships.

A similar notion is behind one of the emerging "hot" topics of computer science, namely *scientific visualization*. The objective of visualization is not to make pretty pictures. It is to render complex data in a visual form that enables us to understand better the underlying phenomena. Notation is a tool of thought, and what these emerging technologies provide is a notational engine *par excellence*. One way to apply our second mirror is to look at the notations used to represent various problem domains, and like a "cognitive anthropologist," attempt to infer what assumptions are implicit in the representation as to how people think, solve problems, etc. In most cases, it is likely that the assumptions thus derived will be a significant distortion of reality.

Scientific visualization is an important attempt to eliminate the distortion; however, science is not the only field of endeavour to which the technology's notational powers can be applied. A good example of using the power of visualization in other domains is the *Information Visualizer* (Card, Robertson & MacKinlay, 1991) which is used to represent data used in the context of office work. An example of the system is shown in Fig. 5.



**Figure 5:** Representing hierarchic data in a "cone tree." From Card, Robertson & MacKinlay (1991).

So far, visualization has been primarily a means of presentation. Data is rendered and displayed, but the degree of interaction is minimal (largely due to the computational overhead of the rendering process). People are not sponges, and do not learn by absorbing information. Rather, they construct knowledge intentionally and actively. Simply put, one learns by experience: exploring and doing. If systems are to reflect better this learning process, then they must, like the *Information Visualizer* be interactive and support active exploration. "Interactive visualization" rather than scientific visualization is the better objective, but there is one thing still missing. From the criteria of our first mirror, there is a problem with focusing just on the visual sense. Both our tactile and aural senses are appropriate channels for representation. Hence, where our path should be heading is towards *Interactive Perceptualization*.

Such a base will lead us to systems that better enable us to explore and understand relationships, test hypotheses and deal with the complexity of our lives.



## 5. Mirror Three: Social context

### *Alone in the corner*

Back in grade-school, if I misbehaved, I was taken out of the group and forced to sit alone, usually facing the wall or corner. Now that I've grown up and have a computer, where do I find myself? - out of the group, sitting alone, usually facing the wall or a corner. The reasons are different, but the punishment is the same.

The point is that the design of the technologies that we use in the workplace have largely ignored the social dynamics of how we work. We face walls because the backs of our machines are so ugly and full of cables, that we hide them. We are anchored to our designated position by the umbilical cord connecting our computer to the wall socket. We sit alone because virtually all of our systems assume that we interact with computers in missionary position: one user and one computer face-to-face - no other position allowed.

### *Instruments of change*

Technologies have had a large impact on how we work, with whom we work and who has what power. That will never change. What can change, however, is who or what is in the driver's seat.

In the past, work has been automated and technologies introduced based on what was possible. If a new technology became available, it was placed in the workplace and the organization adjusted accordingly. Since they were the easiest to program, routine tasks were the first to have technological support.

Of all of the user-related changes emerging today, perhaps the most significant is a change from this approach. Organizations and systems designers are beginning to recognize that rather than have the technology dictate the organizational structure, the organization should dictate the technology. People and how they work are beginning to be recognized as the key to improved productivity.

The importance of this change cannot be overemphasized. No matter how perfectly your icons and menus are designed, or how well a system supports the user in performing their job, if they are doing the wrong job, the system is a failure.

For example, placing computers into patrol cars is intended to help police perform his or her job. But if the technology means that the police now devote more time to apprehending the minor offenders that the technology affords (stolen vehicles, unpaid traffic fines, etc.), the system may be a failure. The courts are clogged with minor offenses, and little has been done to help investigate serious crimes.

### *From birthing rights in central America to PARC*

The past ten years have seen the development of a new profession: applied psychology. Traditionally, psychology has been a discipline which analyzed and tried to understand and explain human behavior. Now, largely due to problems encountered in human-computer interaction, there is a new branch of psychology which attempts to apply this understanding in the context of a design art. The shift is from the descriptive to the prescriptive.

What we are seeing today is a very similar phenomenon in the discipline of socio-anthropology. If we want the society and social structures of work (and play) to drive technology, rather than the reverse, the obvious place to look for expertise is in disciplines such as sociology and anthropology. Like psychology, these are traditionally analytic, not

design disciplines. However, change is coming, and a new discipline is being born: *applied socio-anthropology*.

Hence, we have a new breed of anthropologists like Lucy Suchman and Gitte Jordan (who last studied birthing rites in Central America) stalking the halls of Xerox PARC. There, they study the structure of organizations and work, with the intent of laying the foundation for a design art that takes into account the larger social context. Like psychology, socio-anthropology is becoming a prescriptive as well as analytical science.

#### *Groups and the Instruments*

Perhaps these social concerns are most visible in the rapidly emerging areas of *groupware* and computer supported cooperative work (CSCW). The convergence is from two directions. On the one hand we have the theory growing out of the applied social sciences, and at the other, we have the emergence of important enabling technologies, such as local area networks, new displays technologies and video conferencing.

Telecommunications, video and computer LANS are converging, thereby enabling new forms of collaboration, such as Xerox's *Mediaspace* (Stults, 1986), *EuroPARC* (Buxton & Moran, 1990), and *Ubiquitous Computing* (Weiser, 1991). By integrating a range of technologies, both systems permit a degree of *telepresence* previously impossible.

#### *Telepresence*

By "telepresence," we mean the use of technology to support a sense of social proximity, despite geographical and/or temporal distance. Through telepresence, workers who are at remote sites will, ideally, be able to share a degree of collegiality and interaction that would otherwise only have been possible by being co-resident at the same site. Likewise, telepresence technologies will permit those who missed key meetings or classes to attend "after the fact." Thereby, they will be able to capture not only the content of what passed, but also the shared references and organizational culture from which that content grew. As such, telepresence technologies can be thought of as a kind of *social prosthesis* that can help overcome gaps and weaknesses within organizational structures.

Slowly but surely, the emerging technologies are going to let us out of the corner, and permit us to take a full and active role in the group, whether co-located or not.

## **Summary and Conclusions**

The backdrop for this Chapter is the high complexity but limited functionality of current systems. Our strong view is that for systems to reach their full potential, their design must be altered to provide a better match to the skills and potential of their intended users. We have tended too long to try and overcome shortcomings of design through training and documentation. Technology must adapt to the user, not the contrary.

Hopefully, the Three Mirrors model presented in this paper helps provide some catalyst to rethinking how we approach the design of computer systems. Humans have a vast repertoire of skills which they have accumulated from a life time of living in the everyday world. These skills are largely ignored and wasted in current practice. Nevertheless, they are a resource waiting to be exploited by innovative designers. By doing so, designers will greatly improve the fidelity of the reflection in each of our three mirrors, and in so doing,

they will provide us technologies which afford us to do that which we do best, namely, be human.

## Acknowledgments

The work which underlies this chapter has been supported by the Natural and Engineering Sciences Council of Canada, Xerox PARC, Apple Computer Company, Digital Equipment Corp., and the Information Technology Research Centre of Canada. This support is gratefully acknowledged. In addition, I would like to acknowledge the contribution of the Input Research Group of the University of Toronto and my colleague Ron Baecker, who planted the seed for many of my ideas. I would also like to thank Lindsay MacDonald who made many helpful comments on the manuscript. Finally, I would like to thank the Friend21 workshop in Japan which was the initial catalyst for this chapter.

## References

- Brooks, F.P. Jr., Ouh-Young, M, Batter, J. & Kilpatrick, P.J. (1990). Project GROPE - haptic displays for scientific visualization, *Computer Graphics* **24(3)**, Proceedings of SIGGRAPH '90, 177-185.
- Buxton, W. (1986) There's More to Interaction than Meets the Eye: Some Issues in Manual Input. In Norman, D. A. and Draper, S. W. (Eds.), (1986), *User Centered System Design: New Perspectives on Human-Computer Interaction*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 319-337.
- Buxton, W. & Moran, T. (1990). EuroPARC's Integrated Interactive Intermedia Facility (iiif): early experience, In S. Gibbs & A.A. Verrijn-Stuart (Eds.). *Multi-user interfaces and applications*, Proceedings of the IFIP WG 8.4 Conference on Multi-user Interfaces and Applications, Heraklion, Crete. Amsterdam: Elsevier Science Publishers B.V. (North-Holland), pp. 11-34.
- Gaver, W.W. (1989). The SonicFinder: An interface that uses auditory icons. *Human-Computer Interaction* **4(1)**, 67-94.
- Iwata, H. (1990). Artificial reality with force-feedback: development of desktop virtual space with compact master manipulator, *Computer Graphics* **24(3)**, Proceedings of SIGGRAPH '90, 165-170.
- Stults, R. (1986). *Media Space*. Systems Concepts Lab Technical Report. Palo Alto, CA: Xerox PARC.
- Tufte, E. (1983). *The visual display of quantitative information*. Cheshire, Connecticut: Graphics Press.
- Tufte, E. (1990). *Envisioning information*. Cheshire, Connecticut: Graphics Press.
- Weiser, M. (1991). The computer for the 21st century. *Scientific American*, **265(3)**, 94-104.