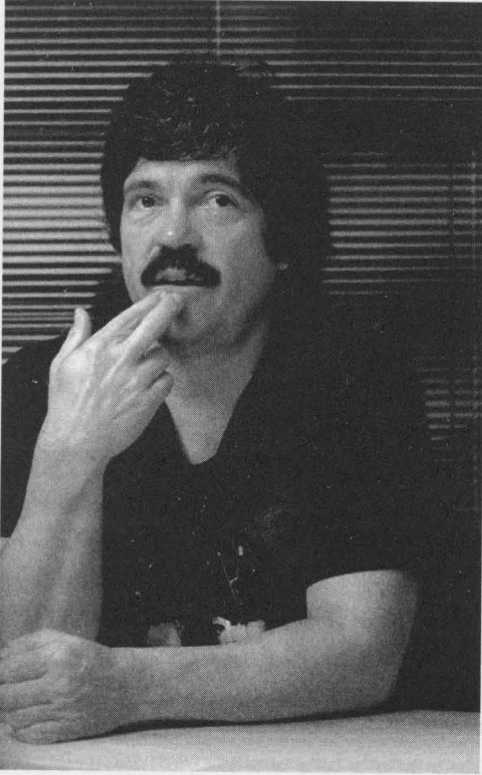Alan Kay

"User Interface: A Personal View" (1989)

Alan Kay. Courtesy of Alan Kay.

"*Therefore, let me argue that the actual dawn of user interface design first happened when computer designers finally noticed, not just that end users had functioning minds, but that a better understanding of how those minds worked would completely shift the paradigm of interaction.*"

<< During the late 1960s, while a graduate student at the University of Utah, Alan Kay studied with Ivan Sutherland, the pioneering scientist whose work launched the field of interactive computer graphics. In 1968 he attended Douglas Engelbart's historic presentation of the oNLine System, which introduced the mouse to computing, suggesting for the first time the possibility of "navigating" information space. In the early 1970s, after forming the Learning Research Group at the newly founded Xerox PARC (Palo Alto Research Center), Kay synthesized these influences into what is considered the most crucial advancement of human-computer interactivity, the graphical user interface (GUI). Kay developed the idea of iconic, graphical representations of computing functions—the folders, menus, and overlapping windows found on the desktop—based on his research into the intuitive processes of learning and creativity.

The GUI was an entirely new approach to interactive computing. The fundamentals of the computer interface had received scant attention since the birth of the first digital computer, ENIAC, twenty-five years before. At the time, computers were programmed and operated almost exclusively by scientists. Kay realized early on that designing an intuitive, easy-to-use interface required a subtle understanding of the dynamics of human perception. Norbert Wiener's concept of the digital computer had been modeled after the brain, with its sophisticated network of message passing and synoptic relays. Kay went deeper in this direction. He drew on the theories of psychologists Jean Piaget, Seymour Papert, and Jerome Bruner, who had studied the intuitive capacities for learning present in the child's mind, and the role that images and symbols play in the building of complex concepts. Kay came to understand, as he put it, that "doing with images makes symbols." This was the premise behind the GUI, which enabled computers users to formulate ideas in real time by manipulating icons on the computer screen.

Kay's approach made computers accessible to nonspecialists. More important, it transformed the computer into a vehicle for popular creative expression. In this essay, Kay describes his realization that the computer was not just a computational device but a medium in its own right. Computers, Kay recognized, might one day replace books. This led him to build a prototype for a personal computer, the Dynabook, which would be capable of dynamically representing images and concepts. This notebook-sized computer, able to provide unprecedented access to information, images, sound, and animation, was a revolutionary departure not only from the stasis of the book but also from the passivity of television. While the Dynabook was never built, it led to the Xerox Alto, the first true multimedia machine. >>

When I was asked to write this chapter, my first reaction was "A book on user interface design—does that mean it's now a real subject?" Well, as of 1989, it's still yes and no. User interface has certainly been a hot topic for discussion since the advent of the Macintosh. Everyone seems to want user interface but they are not sure whether they should order it by the yard or by the ton. Many are just now discovering that user interface is not a sandwich spread—applying the Macintosh style to poorly designed applications and machines is like trying to put Béarnaise sauce on a hotdog!

Of course the practice of user interface design has been around at least since humans invented tools. The unknown designer who first put a haft on a hand axe was trying not just to increase leverage but also to make it an extension of the *arm*, not just the fist. The evolutionary designer whom Richard Dawkins calls the Blind Watchmaker has been at it much longer; all of life's startling interfitness is the result. A more recent byproduct of the industrial revolution called *ergonomics* in Europe and *human factors* in the U.S. has studied how the human body uses senses and limbs to work with tools. From the earliest use of interactive computing in the fifties—mostly for air traffic control and defense—there have been attempts at user interface design and application of ergonomic principles. Many familiar components of modern user interface design appeared in the fifties and early sixties, including pointing devices, windows, menus, icons, gesture recognition, hypermedia, the first personal computer, and more. There was even a beautifully designed user interface for an end-user system in JOSS—but its significance was appreciated only by its designer and users.

Therefore, let me argue that the actual dawn of user interface design first happened when computer designers finally noticed, not just that end users had functioning minds, but that a better understanding of how those minds worked would completely shift the paradigm of interaction.

This enormous change in point of view happened to many computerists in the late sixties, especially in the ARPA research community. Everyone had their own catalyst. For me it was the FLEX machine, an early desktop personal computer of the late sixties designed by Ed Cheadle and myself.

Based on much previous work by others, it had a tablet as a pointing device, a high-resolution display for text and animated graphics, and multiple windows, and it directly executed a high-level object-oriented end-user simulation language. And of course it had a "user interface," but one that repelled end users instead of drawing them closer to the hearth. I recently revisited the FLEX machine

design and was surprised to find how modern its components were—even a use of icon-like structures to access previous work.

But the combination of ingredients didn't gel. It was like trying to bake a pie from random ingredients in a kitchen: baloney instead of apples, ground-up Cheerios instead of flour, etc.

Then, starting in the summer of 1968, I got hit on the head randomly but repeatedly by some really nifty work. The first was just a little piece of glass at the University of Illinois. But the glass had tiny glowing dots that showed text characters. It was the first flat-screen display. I and several other grad students wondered when the surface could become large and inexpensive enough to be a useful display. We also wondered when the FLEX machine silicon could become small enough to fit on the back of the display. The answer to both seemed to be the late seventies or early eighties. Then we could all have an inexpensive powerful notebook computer—I called it a "personal computer" then, but I was thinking *intimacy*.

I read McLuhan's *Understanding Media* [1964] and understood that the most important thing about any communications medium is that message receipt is really message recovery; anyone who wishes to receive a message embedded in a medium must first have internalized the medium so it can be "subtracted" out to leave the message behind. When he said "the medium is the message" he meant that you have to *become* the medium if you use it.

That's pretty scary. It means that even though humans are the animals that shape tools, it is in the nature of tools and man that learning to use tools reshapes us. So the "message" of the printed book is, first, its availability to individuals, hence, its potential detachment from extant social processes; second, the uniformity, even coldness, of noniconic type, which detaches readers from the vividness of the now and the slavery of commonsense thought to propel them into a far more abstract realm in which ideas that don't have easy visualizations can be treated.

McLuhan's claim that the printing press was the dominant force that transformed the hermeneutic Middle Ages into our scientific society should not be taken too lightly—especially because the main point is that the press didn't do it just by making books more available, it did it by changing the thought patterns of those who learned to read.

Though much of what McLuhan wrote was obscure and arguable, the sum total to me was a shock that reverberates even now. The computer is a medium! I had always thought of it as a tool, perhaps a vehicle—a much weaker conception. What McLuhan was saying is that if the personal computer is a truly new

medium then the very use of it would actually change the thought patterns of an entire civilization. He had certainly been right about the effects of the electronic stained-glass window that was television—a remedievalizing tribal influence at best. The intensely interactive and involving nature of the personal computer seemed an antiparticle that could annihilate the passive boredom invoked by television. But it also promised to surpass the book to bring about a new kind of renaissance by going beyond static representations to dynamic simulation. What kind of a thinker would you become if you grew up with an active simulator connected, not just to one point of view, but to all the points of view of the ages represented so they could be dynamically tried out and compared? I named the notebook-sized computer idea the Dynabook to capture McLuhan's metaphor in the silicon to come.

Shortly after reading McLuhan, I visited Wally Feurzeig, Seymour Papert, and Cynthia Solomon at one of the earliest LOGO tests within a school. I was amazed to see children writing programs (often recursive) that generated poetry, created arithmetic environments, and translated English into Pig Latin. And they were just starting to work with the new wastepaper-basket-sized turtle that roamed over sheets of butcher paper making drawings with its pen.

I was possessed by the analogy between print literacy and LOGO. While designing the FLEX machine I had believed that end users needed to be able to program before the computer could become truly theirs—but here was a real demonstration, and with children! The ability to "read" a medium means you can *access* materials and tools created by others. The ability to "write" in a medium means you can *generate* materials and tools for others. You must have both to be literate. In print writing, the tools you generate are rhetorical; they demonstrate and convince. In computer writing, the tools you generate are processes; they simulate and decide.

If the computer is only a vehicle, perhaps you can wait until high school to give "driver's ed" on it—but if it's a medium, then it must be extended all the way into the world of the child. How to do it? Of course it has to be done on the intimate notebook-sized Dynabook! But how would anyone "read" the Dynabook, let alone "write" on it?

LOGO showed that a special language designed with the end user's characteristics in mind could be more successful than a random hack. How had Papert learned about the nature of children's thought? From Jean Piaget, the doyen of European cognitive psychologists. One of his most important contributions is the idea that children go through several distinctive intellectual stages as they develop from birth to maturity. Much can be accomplished if the nature of the stages is

heeded and much grief to the child can be caused if the stages are ignored. Piaget noticed a kinesthetic stage, a visual stage, and a symbolic stage. An example is that children in the visual stage, when shown a squat glass of water poured into a tall thin one, will say there is more water in the tall thin one even though the pouring was done in front of their eyes.

One of the ways Papert used Piaget's ideas was to realize that young children are not well equiped to do "standard" symbolic mathematics until the age of 11 or 12, but that even very young children can do other kinds of math, even advanced math such as topology and differential geometry, when it is presented in a form that is well matched to their current thinking processes. The LOGO turtle with its local coordinate system (like the child, it is always at the center of its universe) became a highly successful "microworld" for exploring ideas in differential geometry.

This approach made a big impression on me and got me to read many more psychology books. Most (including Piaget's) were not very useful, but then I discovered Jerome Bruner's *Towards a Theory of Instruction* [1966]. He had repeated and verified many of Piaget's results, and in the process came up with a different and much more powerful way to interpret what Piaget had seen. For example, in the water-pouring experiment, after the child asserted there was more water in the tall thin glass, Bruner covered it up with a card and asked again. This time the child said, "There must be the same because where would the water go?" When Bruner took away the card to again reveal the tall thin glass, the child immediately changed back to saying there was more water.

When the cardboard was again interposed the child changed yet again. It was as though one set of processes was doing the reasoning when the child could see the situation, and another set was employed when the child could not see. Bruner's interpretation of experiments like these is one of the most important foundations for human-related design. Our mentalium seems to be made up of multiple separate mentalities with very different characteristics. They reason differently, have different skills, and often are in conflict. Bruner identified a separate mentality with each of Piaget's stages: he called them *enactive, iconic, symbolic*. While not ignoring the existence of other mentalities, he concentrated on these three to come up with what are still some of the strongest ideas for creating learning-rich environments.

The work of Papert convinced me that whatever user interface design might be, it was solidly intertwined with learning. Bruner convinced me that learning takes place best environmentally and roughly in stage order—it is best to learn something kinesthetically, then iconically, and finally the intuitive knowledge will

be in place that will allow the more powerful but less vivid symbolic processes to work at their strongest. This led me over the years to the pioneers of environmental learning: Montessori Method, Suzuki Violin, and Tim Gallwey's *Inner Game of Tennis,* to name just a few.

My point here is that as soon as I was ready to look deeply at the human element, and especially after being convinced that the heart of the matter lay with Bruner's multiple mentality model, I found the knowledge landscape positively festooned with already accomplished useful work. It was like the man in Moliere's *Bourgeois Gentilhomme* who discovered that all his life he had been speaking prose! I suddenly remembered McLuhan: "I don't know who discovered water but it wasn't a fish." Because it is in part the duty of consciousness to represent ourselves to ourselves as simply as possible, we should sorely distrust our commonsense self view. It is likely that this mirrors-within-mirrors problem in which we run into a misleading commonsense notion about ourselves at every turn is what forced psychology to be one of the most recent sciences—if indeed it yet is.

Now, if we agree with the evidence that the human cognitive facilities are made up of a *doing* mentality, an *image* mentality, and a *symbolic* mentality, then any user interface we construct should at least cater to the mechanisms that seem to be there. But how? One approach is to realize that no single mentality offers a complete answer to the entire range of thinking and problem solving. User interface design should integrate them at least as well as Bruner did in his spiral curriculum ideas.

One of the implications of the Piaget-Bruner decomposition is that the mentalities originated at very different evolutionary times and there is little probability that they can intercommunicate and synergize in more than the most rudimentary fashion. In fact, the mentalities are more likely to interfere with each other as they compete for control. The study by Hadamard on math and science creativity [1945] and others on music and the arts indicate strongly that creativity in these areas is not at all linked to the symbolic mentality (as most theories of teaching suppose), but that the important work in creative areas is done in the initial two mentalities—most in the iconic (or figurative) and quite a bit in the enactive. The groundbreaking work by Tim Gallwey on the "inner game" [1974] showed what could be done if interference were removed (mentalities not relevant to the learning were distracted) and attention was facilitated (the mentalities that could actually do the learning were focused more strongly on the environment).

Finally, in the sixties a number of studies showed just how modeful was a

mentality that had "seized control"—particularly the analytical-problem-solving one (which identifies most strongly with the Bruner symbolic mentality). For example, after working on five analytic tasks in a row, if a problem was given that was trivial to solve figuratively, the solver could be blocked for hours trying to solve it symbolically. This makes quite a bit of sense when you consider that the main jobs of the three mentalities are:

enactive   know where you are, manipulate
iconic   recognize, compare, configure, concrete
symbolic   tie together long chains of reasoning, abstract

The visual system's main job is to be interested in everything in a scene, to dart over it as one does with a bulletin board, to change context. The symbolic system's main job is to stay with a context and to make indirect connections. Imagine what it would be like if it were reversed. If the visual system looked at the object it first saw in the morning for five hours straight! Or if the symbolic system couldn't hold a thought for more than a few seconds at a time!

It is easy to see that one of the main reasons that the figurative system is so creative is that it tends not to get blocked because of the constant flitting and darting. The chance of finding an interesting pattern is very high. It is not surprising, either, that many people who are "figurative" have extreme difficulty getting anything finished—there is always something new and interesting that pops up to distract. Conversely, the "symbolic" person is good at getting things done, because of the long focus on single contexts, but has a hard time being creative, or even being a good problem solver, because of the extreme tendency to get blocked. In other words, because none of the mentalities is supremely useful to the exclusion of the others, the best strategy would be to try to gently force synergy between them in the user interface design.

Out of all this came the main slogan I coined to express this goal:

Doing with Images makes Symbols

The slogan also implies—as did Bruner—that one should start with—be grounded in—the concrete "Doing with Images," and be carried into the more abstract "makes Symbols."

All the ingredients were already around. We were ready to notice what the theoretical frameworks from other fields of Bruner, Gallwey, and others were trying to tell us. What is surprising to me is just how long it took to put it all together. After Xerox PARC provided the opportunity to turn these ideas into reality, it still took our group about five years and experiments with hundreds of users to come

up with the first practical design that was in accord with Bruner's model and really worked.

DOING	mouse	*enactive*	know where you are, manipulate
with IMAGES	icons, windows	*iconic*	recognize, compare, configure, concrete
makes SYMBOLS	Smalltalk	*symbolic*	tie together long chains of reasoning, abstract

Part of the reason perhaps was that the theory was much better at confirming that an idea was good than at actually generating the ideas. In fact, in certain areas like "iconic programming," it actually held back progress, for example, the simple use of icons as signs, because the siren's song of trying to do symbolic thinking iconically was just too strong.

Some of the smaller areas were obvious and found their place in the framework immediately. Probably the most intuitive was the idea of multiple overlapping windows. NLS had multiple panes, FLEX had multiple windows, and the bit-map display that we thought was too small, but that was made from individual pixels, led quickly to the idea that windows could appear to overlap. The contrastive ideas of Bruner suggested that there should always be a way to compare. The flitting-about nature of the iconic mentality suggested that having as many resources showing on the screen as possible would be a good way to encourage creativity and problem solving and prevent blockage. An *intuitive* way to use the windows was to activate the window that the mouse was in and bring it to the "top." This interaction was *modeless* in a special sense of the word. The active window constituted a mode to be sure—one window might hold a painting kit, another might hold text—but one could get to the next window to do something in *without any special termination*. This is what *modeless* came to mean for me— the user could always get to the next thing desired without any backing out. The contrast of the nice modeless interactions of windows with the clumsy command syntax of most previous systems directly suggested that everything should be made modeless. Thus began a campaign to "get rid of modes."

The object-oriented nature of Smalltalk was very suggestive. For example, *object-oriented* means that the object knows what it can do. In the abstract symbolic arena, it means we should first write the object's name (or whatever will

fetch it) and then follow with a message it can understand that asks it to do some-thing. In the concrete user-interface arena, it suggests that we should select the object first. It can then furnish us with a menu of what it is willing to do. In both cases we have the *object* first and the *desire* second. This unifies the concrete with the abstract in a highly satisfying way.

The most difficult area to get to be modeless was a very tiny one, that of ele-mentary text editing. How to get rid of "insert" and "replace" modes that had plagued a decade of editors? Several people arrived at the solution simultane-ously. My route came as the result of several beginning programmer adults who were having trouble building a paragraph editor in Smalltalk, a problem I thought should be easy. Over a weekend I built a sample paragraph editor whose main simplification was that it eliminated the distinction between insert, replace, and delete by allowing selections to extend *between* the characters. Thus, there could be a zero-width selection, and thus every operation could be a replace. "Insert" meant replace the zero-width selection. "Delete" meant replace the selection with a zero-width string of characters. I got the tiny one-page program running in Smalltalk and came in crowing over the victory. Larry Tesler thought it was great and showed me the idea, already working in his new Gypsy editor (which he implemented on the basis of a suggestion from Peter Deutsch). So much for creativity and invention when ideas are in the air. As Goethe noted, the most im-portant thing is to enjoy the thrill of discovery rather than to make vain attempts to claim priority! . . .

The only stumbling place for this onrushing braver new world is that all of its marvels will be very difficult to communicate with, because, as always, the user interface design that could make it all simple lags far, far behind. If *communica-tion* is the watchword, then what do we communicate with and how do we do it? We communicate with:

- Our selves
- Our tools
- Our colleagues and others
- Our agents

Until now, personal computing has concentrated mostly on the first two. Let us now extend everything we do to be part of a *grand collaboration*—with one's self, one's tools, other humans, and, increasingly, with *agents:* computer processes that act as guide, as coach, and as amanuensis. The user interface design will be the critical factor in the success of this new way to work and play on the com-

puter. One of the implications is that the "network" will not be seen at all, but rather "felt" as a shift in capacity and range from that experienced via one's own hard disk. . . .

Well, there are so many more new issues that must be explored as well. I say thank goodness for that. How do we navigate in once-again uncharted waters? I have always believed that of all the ways to approach the future, the vehicle that gets you to the most interesting places is Romance. The notion of tool has always been a romantic idea to humankind—from swords to musical instruments to personal computers, it has been easy to say: "The best way to predict the future is to invent it!" The romantic dream of "how nice it would be if . . ." often has the power to bring the vision to life. Though the notion of management of complex processes has less cachet than that of the hero singlehandedly wielding a sword, the real romance of management is nothing less than the creation of civilization itself. What a strange and interesting frontier to investigate. As always, the strongest weapon we have to explore this new world is the one between our ears—providing it's loaded!

multiMEDIA
FROM WAGNER TO VIRTUAL REALITY

Edited by
RANDALL PACKER
KEN JORDAN

Foreword by
WILLIAM GIBSON

Coda by
LAURIE ANDERSON

"For anyone who wants to know where multimedia
technology is going, or where it has been."
—BOSTON GLOBE