

USER-CENTERED DESIGN

!*~!*

OFF THE LEASH By W.B. Park



"Darn these hooves! I hit the wrong switch again! Who designs these instrument panels, raccoons?"

The point of POET is to advocate a user-centered design, a philosophy based on the needs and interests of the user, with an emphasis on making products usable and understandable. In this chapter I summarize the main principles, discuss some implications, and offer suggestions for the design of everyday things.

Design should:

- Make it easy to determine what actions are possible at any moment (make use of constraints).
- Make things visible, including the conceptual model of the system, the alternative actions, and the results of actions.
- Make it easy to evaluate the current state of the system.
- Follow natural mappings between intentions and the required actions; between actions and the resulting effect; and between the information that is visible and the interpretation of the system state.

In other words, make sure that (1) the user can figure out what to do, and (2) the user can tell what is going on.

Design should make use of the natural properties of people and of the world: it should exploit natural relationships and natural constraints. As much as possible, it should operate without instructions or labels. Any necessary instruction or training should be needed only once; with each explanation the person should be able to say, "Of course," or "Yes, I see." A simple explanation will suffice if there is reason to the design, if everything has its place and its function, and if the outcomes of actions are visible. If the explanation leads the person to think or say, "How am I going to remember that?" the design has failed.

Seven Principles for Transforming Difficult Tasks into Simple Ones

How does the designer go about the task? As I've argued in POET, the principles of design are straightforward.

1. Use both knowledge in the world and knowledge in the head.
2. Simplify the structure of tasks.
3. Make things visible: bridge the gulfs of Execution and Evaluation.
4. Get the mappings right.

5. Exploit the power of constraints, both natural and artificial.
6. Design for error.
7. When all else fails, standardize.

USE BOTH KNOWLEDGE IN THE WORLD AND KNOWLEDGE IN THE HEAD

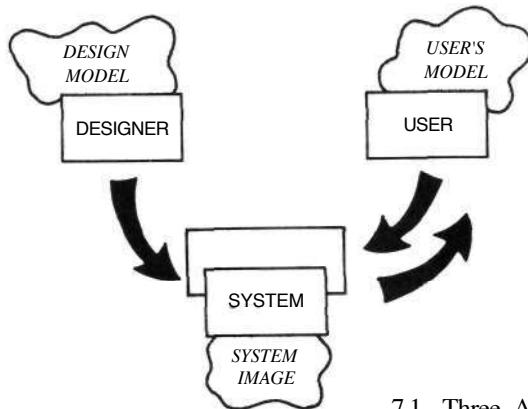
I have argued that people learn better and feel more comfortable when the knowledge required for a task is available externally—either explicit in the world or readily derived through constraints. But knowledge in the world is useful only if there is a natural, easily interpreted relationship between that knowledge and the information it is intended to convey about possible actions and outcomes.

Note, however, that when a user is able to internalize the required knowledge—that is, to get it into the head—performance can be faster and more efficient. Therefore, the design should not impede action, especially for those well-practiced, experienced users who have internalized the knowledge. It should be easy to go back and forth, to combine the knowledge in the head with that in the world. Let whichever is more readily available at the moment be used without interfering with the other, and allow for mutual support.

THREE CONCEPTUAL MODELS

The operation of any device—whether it be a can opener, a power generating plant, or a computer system—is learned more readily, and the problems are tracked down more accurately and easily, if the user has a good conceptual model. This requires that the principles of operation be observable, that all actions be consistent with the conceptual model, and that the visible parts of the device reflect the current state of the device in a way consistent with that model. The designer must develop a conceptual model that is appropriate for the user, that captures the important parts of the operation of the device, and that is understandable by the user.

Three different aspects of mental models must be distinguished: the *design model*, the *user's model*, and the *system image* (figure 7.1). The design model is the conceptualization that the designer has in mind. The user's model is what the user develops to explain the operation of the system. Ideally, the user's model and the design model are equivalent. How-



7.1 Three Aspects of Mental Models.
The design model, the user's model, and the system image. (From Norman, 1986.)

ever, the user and designer communicate only through the system itself: its physical appearance, its operation, the way it responds, and the manuals and instructions that accompany it. Thus the *system image* is critical: the designer must ensure that everything about the product is consistent with and exemplifies the operation of the proper conceptual model.

All three aspects are important. The user's model is essential, of course, for that determines what is understood. In turn, it is up to the designer to start with a design model that is functional, learnable, and usable. The designer must ensure that the system reveals the appropriate system image. Only then can the user acquire the proper user's model and find support for the translation of intentions into actions and system state into interpretations. Remember, the user acquires all knowledge of the system from that system image.

THE ROLE OF MANUALS

The system image includes instruction manuals and documentation.

Manuals tend to be less helpful than they should be. They are often written hastily, after the product is designed, under severe time pressures and with insufficient resources, and by people who are overworked and underappreciated. In the best of worlds, the manuals would be written first, then the design would follow the manual. While the product was being designed, potential users could simultaneously

test the manuals and mock-ups of the system, giving important design feedback about both.

Alas, even the best manuals cannot be counted on; many users do not read them. Obviously it is wrong to expect to operate complex devices without instruction of some sort, but the designers of complex devices have to deal with human nature as it is.

SIMPLIFY THE STRUCTURE OF TASKS

Tasks should be simple in structure, minimizing the amount of planning or problem solving they require. Unnecessarily complex tasks can be restructured, usually by using technological innovations.

Here is where the designer must pay attention to the psychology of the person, to the limits on how much a person can hold in memory at one time, to the limits on how many active thoughts can be pursued at once. These are the limitations of short-term and long-term memory and of attention. The limitations of short-term memory (STM) are such that a person should not be required to remember more than about five unrelated items at one time. If necessary, the system should provide technological assistance for any temporary memory requirements. The limitations of long-term memory (LTM) mean that information is better and more easily acquired if it makes sense, if it can be integrated into some conceptual framework. Moreover, retrieval from LTM is apt to be slow and to contain errors. Here is where information in the world is important, to remind us of what can be done and how to do it. Limitations on attention are also severe; the system should help by minimizing interruption, by providing aids to allow for recovery of the exact status of the operations that were interrupted.

A major role of new technology should be to make tasks simpler. A task can be restructured through technology, or technology might provide aids to reduce the mental load. Technological aids can show the alternative courses of action; help evaluate implications; and portray outcomes in a more complete, more easily interpretable manner. These aids can make the mappings more visible or, better, make the mappings more natural. Four major technological approaches can be followed:

- Keep the task much the same, but provide mental aids.
- Use technology to make visible what would otherwise be invisible, thus improving feedback and the ability to keep control.

- Automate, but keep the task much the same.
- Change the nature of the task.

Let us look separately at each of these possibilities.

KEEP THE TASK MUCH THE SAME, BUT PROVIDE MENTAL AIDS

Don't underestimate the power or importance of simple mental aids. Consider, for example, the value of simple, everyday notes to ourselves. Without them, we might fail. Or simple notepads for telephone numbers, names, addresses—for the facts that are essential to everyday functioning, but that we cannot trust our own memory structures to provide. Some mental aids are also technological advances; these include watches, timers, calculators, pocket dictating machines, computer notepads, and computer alarms. Some aids are still to come: the pocket computer with a powerful display, which will keep our notes, remind us of our appointments, and smooth our passage through the schedules and interactions of life.

USE TECHNOLOGY TO MAKE VISIBLE WHAT WOULD OTHERWISE BE INVISIBLE, THUS IMPROVING FEEDBACK AND THE ABILITY TO KEEP CONTROL

The instruments in the automobile or aircraft do not change the task, but they do make visible the state of the engine and the other parts of the vehicle, even though you cannot physically get access to them. Similarly, the microscope and telescope, television set, camera, microphone, and loudspeaker all provide ways of getting information about a remote object, making visible (or audible) what is happening, making possible tasks and pursuits that would otherwise not be possible. With modern computers and their powerful graphic displays, we now have the power to show what is really happening, to provide a good, complete image that matches the person's mental model of the task—thereby simplifying both understanding and performance. Today, computer graphics are used more for show than for legitimate purposes. Their powers are wasted. But there exists great potential to make visible what should be visible (and to keep hidden what is irrelevant).

These first two approaches to mental aids keep the main tasks unchanged. They act as reminders. They reduce memory load by provid-

ing external memory devices (providing knowledge in the world rather than requiring it to be in the head). They supplement our perceptual abilities. Sometimes they enhance human skills sufficiently so that a job that was not possible before, or was possible only for the most highly skilled performers, becomes available to many.

Don't these so-called advances also cause us to lose valuable mental skills? Each technological advance that provides a mental aid also brings along critics who decry the loss of the human skill that has been made less valuable. Fine, I say: if the skill is easily automated, it wasn't essential.

I prefer to remember things by writing them on a pad of paper rather than spending hours of study on the art of memory. I prefer using a pocket calculator to spending hours of pencil pushing and grinding, usually only to make an arithmetic mistake and not discover it until after the harm has been done. I prefer prerecorded music to no music, even if I risk becoming complacent about the power and beauty of the rare performance. And I prefer writing on a text editor or word processor so that I can concentrate on the ideas and the style, not on making marks on the paper. Then I can go back later and correct ideas, redo the grammar. And with the aid of my all-important spelling correction program, I can be confident of my presentation.

Do I fear that I will lose my ability to spell as a result of overreliance on this technological crutch? What ability? Actually, my spelling is improving through the use of this spelling corrector that continually points out my errors and suggests the correction, but won't make a change unless I approve. It is certainly a lot more patient than my teachers used to be. And it is always there when I need it, day or night. So I get continual feedback about my errors, plus useful advice. My typing does seem to be deteriorating because I can now type even more sloppily, confident that my mistakes will be detected and corrected.

In general, I welcome any technological advance that reduces my need for mental work but still gives me the control and enjoyment of the task. That way I can exert my mental efforts on the core of the task, the thing to be remembered, the purpose of the arithmetic or the music. I want to use my mental powers for the important things, not fritter them away on the mechanics.

AUTOMATE, BUT KEEP THE TASK MUCH THE SAME

There are dangers in simplification: unless we are careful, the automation can harm as well as help. Consider one impact of automation.

As before, the task will stay essentially the same, but parts of it will disappear. In some cases the change is confirmed as a universal blessing. I don't know of anyone who misses the automatic spark advance in automobiles or cranking the engine to get it started. Just a few people miss having manual control over the automobile choke. On the whole, this type of automation has resulted in useful advances, replacing tedious or unnecessary tasks and reducing what must be monitored. The automatic controls and instruments of ships and aircraft have been great improvements. Some automation is more problematic. Automatic shift on a car: Do we lose some control, or does it help lighten the mental burden of driving? After all, we drive to get to a destination, so the need to monitor engine speed and gearshift position would seem quite irrelevant. But some people take pleasure in performing the task itself; for them, part of driving is using the engine well, believing that they can operate more efficiently than can the automatic device.

What about the automatic pilot of an aircraft, or the automatic navigation systems that have eliminated the sextant and lengthy computations? Or what about frozen, precooked meals? Do the changes destroy the essence of the task? Here there's more debate. In the best of worlds we would be able to choose automation or full control.

CHANGE THE NATURE OF THE TASK

When a task seems inherently complex because of the manual skill required, certain technological aids can dramatically change which type of skill is required by restructuring the task. In general, technology can help transform deep, wide structures into narrower, shallower ones.

Tying a shoelace is one of the standard, everyday tasks that is actually quite difficult to learn. Adults may have forgotten how long it took them to learn (but they will be reminded if their fingers stiffen with injury, age, or disease). The introduction of new fastening materials—for example, Velcro hook-and-loop fasteners—has eliminated the need for a complex sequence of skilled motor actions by changing the task to one that is considerably simpler, one that requires less skill. The task has become possible for both young children and infirm adults. The example of shoelaces may seem trivial, but it isn't; like many everyday activities, it is difficult for a large segment of the population and its difficulties can be overcome through the restructuring provided by a simple technology.

The hook-and-loop fasteners provide another example of design tradeoffs (figure 7.2). Hook-and-loop fasteners dramatically simplify shoe fastening for the young and infirm. But they add to the problems of parents and teachers, for children delight in fastening and unfastening their shoes; so a fastener that is more difficult to work has certain virtues. And for sports for which precise support of the foot is required, the best solution still appears to be the shoelace, which can be adjusted so as to offer different tensions at different parts of the foot. The current generation of hook-and-loop fasteners does not have the flexibility of laces.

Digital watches represent another example of how a new technology can supplant an old one; it has delayed or eliminated the need for children to learn the mapping of the analog hands of the traditional clockface onto the hours, minutes, and seconds of the day. Digital timepieces are controversial: in changing the representation of time, the

7.2 Hook-and-Loop Fastener. With the use of hook-and-loop fasteners, the act of tying shoes is much simplified: a good example of the power of technology to change the nature of the task. But there is a cost. Children find the task so easy they gleefully untie their shoes. And these fasteners are not yet as flexible as shoelaces for the support needed for sports.



power of the analog form has been lost, and it has become more difficult to make quick judgments about time. The digital display makes it easier to determine the exact time, but harder to make estimates or to see approximately how much time has passed since an earlier reading. This might serve as a useful reminder that task simplification, by itself, is not necessarily a virtue.

I do not want to argue for digital timepieces, but let me remind you how difficult and arbitrary the analog timepiece really is. After all, it, too, was an arbitrary imposition of a notational scheme, imposed upon the world by the early technologists. Today, because we can no longer remember the origins, we think of the analog system as necessary, virtuous, and proper. It presents a horrid, classic example of the mapping problem. Yes, the notion that time should be represented by the distance a hand moves around a circle is a good one. The problem is that we use two or three different hands moving around the same circle, each one meaning something different and operating on a different scale. Which hand is which? (Do you remember how hard it is to teach a child the difference between the little hand and the big hand, and not to confuse the second hand—which is sometimes big, sometimes little—with the minute hand or the hour hand?)

*Do I exaggerate? Read what Kevin Lynch says about this in his delightful book on city planning, *What time is this place?**

"Telling time is a simple technical problem, but unfortunately the clock is a rather obscure perceptual device. Its first widespread use in the thirteenth century was to ring the hours for clerical devotions. The clockface which translated time into spatial alteration, came later. That form was dictated by its works, not by any principle of perception. Two (sometimes three) superimposed cycles give duplicate readings, according to angular displacement around a finely marked rim. Neither minutes nor hours nor half days correspond to the natural cycles of our bodies or the sun. And so teaching a child to read a clock is not a childish undertaking. When asked why a clock had two hands, a four-year-old replied, 'God thought it would be a good idea!'"

Aircraft designers started using meters that looked like clockfaces to represent altitude. As airplanes were able to fly higher and higher, the meters needed more hands. Guess what? Pilots made errors—serious errors. Multihanded analog altimeters have been largely abandoned in favor of digital ones because of the prevalence of reading

errors. Even so, many contemporary altimeters maintain a mixed mode: information about rate and direction of altitude change is determined from a single analog hand, while precise judgments of height come from the digital display,

DON'T TAKE AWAY CONTROL

Automation has its virtues, but automation is dangerous when it takes too much control from the user. "Overautomation"—too great a degree of automation—has become a technical term in the study of automated aircraft and factories.² One problem is that overreliance on automated equipment can eliminate a person's ability to function without it, a prescription for disaster if, for example, one of the highly automated mechanisms of an aircraft suddenly fails. A second problem is that a system may not always do things exactly the way we would like, but we are forced to accept what happens because it is too difficult (or impossible) to change the operation. A third problem is that the person becomes a servant of the system, no longer able to control or influence what is happening. This is the essence of the assembly line: it depersonalizes the job, it takes away control, it provides, at best, a passive or third-person experience.

All tasks have several layers of control. The lowest level is the details of the operation, the nimble finger work of sewing or playing the piano, the nimble mental work of arithmetic. Higher levels of control affect the overall task, the direction in which the work is going. Here we determine, supervise, and control the overall structure and goals. Automation can work at any level. Sometimes we really want to maintain control at the lower level. For some of us, it is the nimble execution of the finger or mind that matters. Some of us want to play music with skill. Or we like the feel of tools against wood. Or we enjoy wielding a paintbrush. In cases like these, we would not want automation to interfere. At other times we want to concentrate on higher level things. Perhaps our goal is to listen to music, and we find the radio more effective for us than the piano; perhaps our artistic skill can't get us as far as can a computer program.

MAKE THINGS VISIBLE: BRIDGE THE GULFS OF EXECUTION AND EVALUATION

This has been a focal theme of POET. Make things visible on the execution side of an action so that people know what is possible and

how actions should be done; make things visible on the evaluation side so that people can tell the effects of their actions.

There is more. The system should provide actions that match intentions. It should provide indications of system state that are readily perceivable and interpretable and that match intentions and expectations. And, of course, the system state should be visible (or audible) and readily interpretable. Make the outcomes of an action obvious.

Sometimes the wrong things are visible. A friend of mine, a professor of computer science at my university, proudly showed me his new CD player and its associated remote control. Sleek, functional. The remote control unit had a little metal loop protruding from one end. When I asked what it was for, my friend told a story. When he first got the set, he assumed that the loop was an antenna for the remote unit, so he always aimed it at the CD player. It didn't seem to work well; he had to stand within a few feet of the CD while using the remote. He mumbled to himself that he had bought a poorly designed unit. Weeks later he discovered that the metal hook was just a hook for hanging up the device. He had been aiming the remote at his own body. When he turned the remote around, it worked from far across the room.

Here is a case of natural mappings that fails. The hook provided a natural mapping for function: it indicated which side of the remote control device should be pointed at the CD set. Unfortunately, it provided erroneous information. In making things visible, it is important to make the correct things visible. Otherwise people form explanations for the things they can see, explanations that are likely to be false. And then they find some reason for poor performance—in this example, that the remote was not very powerful. People are very good at forming explanations, at creating mental models. It is the designer's task to make sure that they form the correct interpretations, the correct mental models: the system image plays the key role.

Remote transmitter units that need to be pointed at a receiver should have some visible evidence of the transmitting mechanism. Modem units carefully hide any indication of the signaling method, violating the rules of visibility. My friend searched hard for some clue of the direction to point the device in, and he found one: the hook. And, no, the instruction manual did not say which end of the unit should be pointed at the CD player.

GET THE MAPPINGS RIGHT

Exploit natural mappings. Make sure that the user can determine the relationships:

- Between intentions and possible actions
- Between actions and their effects on the system
- Between actual system state and what is perceivable by sight, sound, or feel
- Between the perceived system state and the needs, intentions, and expectations of the user

Natural mappings are the basis of what has been called "response compatibility" within the fields of human factors and ergonomics. The major requirement of response compatibility is that the spatial relationship between the positioning of controls and the system or objects upon which they operate should be as direct as possible, with the controls either on the objects themselves or arranged to have an analogical relationship to them. In similar fashion, the movement of the controls should be similar or analogous to the expected operation of the system. Difficulties arise wherever the positioning and movements of the controls deviate from strict proximity, mimicry, or analogy to the things being controlled.

The same arguments apply to the relationship of system output to expectations. A critical part of an action is the evaluation of its effects. This requires timely feedback of the results. The feedback must provide information that matches the user's intentions and must be in a form that is easy to understand. Many systems omit the relevant visible outcomes of actions; even when information about the system state is provided, it may not be easy to interpret. The easiest way to make things understandable is to use graphics or pictures. Modern systems (especially computer systems) are quite capable of this, but the need seems not to have been recognized by designers.

EXPLOIT THE POWER OF CONSTRAINTS, BOTH NATURAL AND ARTIFICIAL

Use constraints so that the user feels as if there is only one possible thing to do—the right thing, of course. In chapter 4 I used the example

of the Lego toy motorcycle, which could be correctly put together by people who had never before seen it. Actually, the toy is not simple. It was carefully designed. It exploits a variety of constraints. It is a good example of the power of natural mappings and constraints, constraints that reduce the number of alternative actions at each step to at most a few.

DESIGN FOR ERROR

Assume that any error that can be made will be made. Plan for it. Think of each action by the user as an attempt to step in the right direction; an error is simply an action that is incompletely or improperly specified. Think of the action as part of a natural, constructive dialog between user and system. Try to support, not fight, the user's responses. Allow the user to recover from errors, to know what was done and what happened, and to reverse any unwanted outcome. Make it easy to reverse operations; make it hard to do irreversible actions. Design explorable systems. Exploit forcing functions.

WHEN ALL ELSE FAILS, STANDARDIZE

When something can't be designed without arbitrary mappings and difficulties, there is one last route: standardize. Standardize the actions, outcomes, layout, displays. Make related actions work in the same way. Standardize the system, the problem; create an international standard. The nice thing about standardization is that no matter how arbitrary the standardized mechanism, it has to be learned only once. People can learn it and use it effectively. This is true of typewriter keyboards, traffic signs and signals, units of measurement, and calendars. When followed consistently, standardization works well.

There are difficulties. It may be hard to obtain an agreement. And timing is crucial: it is important to standardize as soon as possible—to save everyone trouble—but late enough to take into account advanced technologies and procedures. The shortcomings of early standardization are often more than made up for by the increase in ease of use.³

Users have to be trained to the standards. The very conditions that require standardization require training, sometimes extensive training (that is OK: it takes months to learn the alphabet, or to type, or to drive

7.3 The Backward Clock.
(Drawing by Eileen Conway.)



a car). Remember, standardization is essential only when all the necessary information cannot be placed in the world or when natural mappings cannot be exploited. The role of training and practice is to make the mappings and required actions more available to the user, overcoming any shortcomings in the design, minimizing the need for planning and problem solving.

Take the everyday clock. It's standardized. Consider how much trouble you would have telling time with a backward clock, where the hands revolved counterclockwise. Such clocks do exist (figure 7.3). They make effective conversation pieces. Not so good for telling the time, though. Why not? There is nothing illogical about a clock that goes counterclockwise. It's just as logical as one that goes clockwise. The reason we dislike it is that we have standardized on a different scheme, on the very definition of the term "clockwise." Without such standardization, clock reading would be more difficult: you'd always have to figure out the mapping.

STANDARDIZATION AND TECHNOLOGY

If we examine the history of advances in all technological fields, we see that some improvements naturally come through technology, others come through standardization. The early history of the automobile is a good example. The first cars were very difficult to operate. They required strength and skill beyond the abilities of many. Some prob-

lems were solved through automation: the choke, the spark advance, and the starter engine.

Arbitrary aspects of cars and driving had to be standardized:

- Which side of the road people drove on
- Which side of the car the driver sat on
- Where the essential components were: steering wheel, brake, clutch pedal, and accelerator (in some early cars it was on a hand lever)

Standardization is simply another aspect of cultural constraints. With standardization, once you have learned to drive one car, you feel justifiably confident that you can drive any car, any place in the world.

Today's computers are still poorly designed, at least from the user's point of view. But one problem is simply that the technology is still very primitive—like the 1906 auto—and there is no standardization. Standardization is the solution of last resort, an admission that we cannot solve the problems in any other way. So we must at least all agree to a common solution. When we have standardization of our keyboard layouts, our input and output formats, our operating systems, our text editors and word processors, and the basic means of operating any program, then suddenly we will have a major breakthrough in usability.⁴

THE TIMING OF STANDARDIZATION

Standardize and you simplify lives: everyone learns the system only once. But don't standardize too soon; you may be locked into a primitive technology, or you may have introduced rules that turn out to be grossly inefficient, even error-inducing. Standardize too late and there may already be so many ways of doing the task that no international standard can be agreed on; if there is agreement on an old-fashioned technology, it may be too expensive to change. The metric system is a good example: it is a far simpler and more usable scheme for representing distance, weight, volume, and temperature than the older, British system (feet, pounds, seconds, degrees on the Fahrenheit scale). But industrial nations with a heavy commitment to the old measurement standards claim they cannot afford the massive costs and confusion of conversion. So we are stuck with two standards, at least for a few more decades.

Would you consider changing how we specify time? The current system is arbitrary. The day is divided into twenty-four rather arbitrary units—hours. But we tell time in units of twelve, not twenty-four, so there have to be two cycles of twelve hours each, plus the special convention of A.M. and P.M. so we know which cycle we are talking about. Then we divide each hour into sixty minutes and each minute into sixty seconds. What if we switched to metric divisions: seconds divided into tenths, milliseconds, and microseconds? We would have days, millidays, and microdays. There would have to be a newhour, minute, and second: call them the newhour, the newminute, and the newsecond. It would be easy: ten newhours to the day, one hundred newminutes to the newhour, one hundred newseconds to the newminute.

Each newhour would last exactly 2.4 times an old hour: 144 old minutes. So the old one-hour period of the schoolroom or television program would be replaced with a half-newhour period—only 20 percent longer than the old. Each newminute would be quite similar to the current minute: 0.7 of an old minute, to be exact (each newminute would be about 42 old seconds). And each newsecond would be slightly shorter than an old second. The differences in durations could be gotten used to; they aren't that large. And computations would be so much easier. I can hear the everyday conversations now:

"I'll meet you at noon—5 newhours. Don't be late, it's only a half hour from now, 50 newminutes, OK?"

"What time is it? 7.85—15 minutes to the evening news."

"What do I think of it? I wouldn't go near it."

Deliberately Making Things Difficult

"How can good design (design that is usable and understandable) be balanced with the need for 'secrecy' or privacy, or protection? That is, some applications of design involve areas which are sensitive and necessitate strict control over who uses and understands them. Perhaps we don't want any user-in-the-street to understand enough of a system to compromise its security. Couldn't it be argued that some things shouldn't be designed well? Can't things be left cryptic, so that only those who have clearance, extended education, or whatever, can make use of the system? Sure, we have passwords, keys, and other types of



7.4 A School Door, Deliberately Made Difficult to Use. The school is for handicapped children; the school officials did not want children to be able to go in and out of the school without adult supervision. The principles of usability espoused in POET can be followed in reverse to make difficult those tasks that ought to be difficult.

security checks, but this can become wearisome for the privileged user. It appears that if good design is not ignored in some contexts, the purpose for the existence of the system will be nullified. ⁵

Consider figure 7.4, a door on a school in Stapleford, England: the latches are up at the very top of the door, where they are both hard to find and hard to reach. This is *good* design, deliberately and carefully done. The door is to a school for handicapped children, and the school didn't want the children to be able to get out to the street without an adult. Violating the rules of ease of use is just what is needed.

Most things are intended to be easy to use, but aren't. But some things are deliberately difficult to use—and ought to be. The number of things that should be difficult to use is surprisingly large:

- Any door designed to keep people in or out.
- Security systems, designed so that only authorized people will be able to use them.
- Dangerous equipment, which should be restricted.
- Dangerous operations, such as life-threatening actions. These can be designed so that one person alone can't complete the action. I worked for a summer setting off dynamite underwater (to study underwater sound transmission); the circuits were set up to require two people to work them. Two buttons had to be depressed at the same time in order to set off the charge: one button outside, one inside the electronic recording trailer. Similar precautions are taken at military installations.
- Secret doors, cabinets, safes: you don't want the average person even to know that they are there, let alone to be able to work them. These may require two different keys or combinations, meant to be carried or known by two people.
- Cases deliberately intended to disrupt the normal routine action (in chapter 5 I call these forcing functions). Examples include the acknowledgment required before permanently deleting a file from a computer storage system, safeties on pistols and guns, pins in fire extinguishers.
- Controls deliberately made big and spread far apart so that children will have difficulty operating them.
- Cabinets and bottles of medications and dangerous substances deliberately made difficult to open to keep them secure from children.
- Games, a category in which designers deliberately flout the laws of understandability and usability. Games are meant to be difficult. And in some games, such as the adventure or Dungeons and Dragons games popular on home (and office) computers, the whole point of the game is to figure out what is to be done, and how.
- *Not* the door on a train (figure 7.5).

Many things need to be designed for a certain lack of understandability or usability. The rules of design are equally important to know here, however, for two reasons. First, even deliberately difficult designs shouldn't be entirely difficult. Usually there is one difficult part, designed to keep unauthorized people from using the device; the rest of it should follow the normal good principles of design. Second, even if your job is to make something difficult to do, you need to know how

to go about doing it. In this case, the rules are useful, for they state in reverse just how to go about the task. You systematically violate the rules.

- Hide critical components: make things invisible.
- Use unnatural mappings for the execution side of the action cycle, so that the relationship of the controls to the things being controlled is inappropriate or haphazard.
- Make the actions physically difficult to do.
- Require precise timing and physical manipulation.
- Do not give any feedback.
- Use unnatural mappings for the evaluation side of the action cycle, so that system state is difficult to interpret.

Safety systems pose a special problem in design. Oftentimes the design feature added to ensure safety eliminates one danger only to create a secondary one. When workers dig a hole in a street, they must put up barriers to prevent people from walking into the hole. The barriers solve one problem, but they themselves pose another danger, often circumvented by adding signs and flashing lights to warn of the barriers. Emergency doors, lights, and alarms must often be accompanied by warning signs or barriers that control when and how they can be used.

Consider the school door of figure 7.4. Under normal use, this design adds to the safety of the children. But what if there was a fire? Even nonhandicapped adults might have trouble with the door as they rushed to get out. What about short or handicapped teachers—how could they open the door? The solution to one problem—Unauthorized exit of schoolchildren—can easily create a major new problem in times of fire. How could this problem be solved? Probably with a push bar located within everyone's reach on the door, but connected to an alarm so that in normal circumstances it would not be used.

DESIGNING A DUNGEONS AND DRAGONS GAME

One of my students worked for a computer game company helping develop a new Dungeons and Dragons game. He and his fellow students used his experience to do a class project on the difficulty of



To open door:
Lower the window
and use outside handle

Please close window
and shut door after use

7.5 British Train Door, from the Inside. Clearly difficult to use, but why? I haven't the foggiest idea. To prevent accidental opening? To make it so that young children cannot open the door? None of the hypotheses I have tried stand up under close examination. I leave this to the reader.

games. In particular, they combined some research on what makes games interesting with the analysis of the seven stages of action (chapter 2) to determine what factors cause difficulties in dungeon games.⁶ As you might imagine, making things difficult is a tricky business. If a game isn't difficult enough, experienced players lose interest. On the other hand, if it is too difficult, the initial enjoyment gives way to frustration. In fact, several psychological factors hang in a delicate balance: challenge, enjoyment, frustration, and curiosity. As the students reported, "Once the curiosity is lost and the frustration level becomes too high, it is hard to get a person's interest to return to the game." All this has to be considered, yet the game must maintain its appeal for players of many different levels, from first-time players to experienced players. One approach is to sprinkle the game with many different challenges of variable difficulty. Another is to have many little things continually happening, maintaining the curiosity motive.

The same rules that apply to make tasks understandable and usable also apply to make them more difficult and challenging; they can be applied perversely to show where the difficulty should be added. But difficulty and challenge should not be confused with frustration and error. The rules must be applied intelligently, for ease of use or difficulty of use.

EASY LOOKING IS NOT NECESSARILY EASY TO USE

Early in POET I examined the modern office telephone, simple looking but hard to use. I contrasted this with an automobile dashboard that has more than a hundred controls, complicated looking but easy to use. Apparent complexity and actual complexity are not at all the same.

Consider a surfboard, ice skates, parallel bars, or a bugle. All are simple looking. Yet years of study and practice are required to be good at using any of these objects.

The problem is that each of the apparently simple devices is capable of a wide repertoire of actions, but because there are few controls (and no moving parts), the rich complexity of action can be accomplished only through a rich complexity of execution by the user. Remember the office telephone system? When there are more actions than controls, each control must take part in a variety of different actions. If there are exactly the same number of controls as actions, then, in principle, the

controls can be simple and the execution can be simple: find the correct control and activate it.

Actually, increasing the number of controls can both enhance and detract from ease of use. The more controls, the more complex things look and the more the user must learn about; it becomes harder to find the appropriate control at the appropriate time. On the other hand, as the number of controls increases up to the number of functions, there can be a better match between controls and functions, making things easier to use. So the number of controls and complexity of use is really a tradeoff between two opposing factors.

How many controls does a device need? The fewer the controls, the easier it looks to use and the easier it is to find the relevant controls. As the number of controls increases, specific controls can be tailored for specific functions. The device may look more and more complex, but it will be easier to use. We studied this relationship in our laboratory.⁷ Complexity of appearance seems to be determined by the number of controls, whereas difficulty of use is jointly determined by the difficulty of finding the relevant controls (which increases with the number of controls) and difficulty of executing the functions (which may decrease with the number of controls).

We found that to make something easy to use, match the number of controls to the number of functions and organize the panels according to function. To make something look like it is easy, minimize the number of controls. How can these conflicting requirements be met simultaneously? Hide the controls not being used at the moment. By using a panel on which only the relevant controls are visible, you minimize the appearance of complexity. By having a separate control for each function, you minimize complexity of use. It is possible to eat your cake and have it, too.

Design and Society

Tools affect more than the ease with which we do things; they can dramatically affect our view of ourselves, society, and the world. It is hardly necessary to point out the dramatic changes in society that have resulted from the invention of today's everyday things: paper and pencil, the printed book, the typewriter, the automobile, the telephone, radio, and television. Even apparently simple innovations can bring about dramatic changes, most of which cannot be predicted. The tele-

phone, for example, was widely misunderstood ("Why would we want one? Who would we want to talk to?"), as was the computer (fewer than ten *were* thought to be sufficient to satisfy all of America's computing needs).⁸ Predictions of the future of the city were widely off the mark. And nuclear power was once thought destined to lead to atomic automobiles and airplanes. Some people expected private air transportation to become as widespread as the automobile—a helicopter in every garage.

HOW WRITING METHOD AFFECTS STYLE

The history of technology shows that we are not very good at prediction, but that does not diminish the need for sensitivity to possible changes. New concepts will transform society, for better or worse. Let us examine one simple situation: the effect of the gradual automation of the tools of writing on styles of writing.

FROM QUILL AND INK TO KEYBOARD AND MICROPHONE

In earlier times, when goose quill and ink were used on parchment, it was tedious and difficult to correct what had been written. Writers had to be careful. Sentences had to be thought through before being set to paper. One result was sentences that were long and embellished—the graceful rhetorical style we associate with our older literature. With the advent of easier to use writing tools, corrections became easier to make; so writing was done more rapidly, but also with less thought and care—more like everyday speech. Some critics decried the lack of literary niceties. Others argued that this was how people really communicated, and besides, it was easier to understand.

With changes in writing tools, the speed of writing increases. In handwriting, thought runs ahead, posing special demands on memory and encouraging slower, more thoughtful writing. With the typewriter keyboard, the skilled typist can almost keep up with thought. With the advent of dictation, the output and the thought seem reasonably well matched.

Even greater changes have come about with the popularity of dictation. Here the tool can have a dramatic effect, for there is no external record of what has been spoken; the author has to keep everything in memory. As a result, dictated letters often have a long, rambling style.

They are more colloquial and less structured—the former because they are based on speech, the latter because the writer can't easily keep track of what has been said. Style may change further when we get voice typewriters, where our spoken words will appear on the page as they are spoken. This will relieve the memory burden. The colloquial nature may remain and even be enhanced, but—because the printed record of the speech is immediately visible—perhaps the organization will improve.

The widespread availability of computer text editors has produced other changes in writing. On the one hand, it is satisfying to be able to type your thoughts without worrying about minor typographical errors or spelling. On the other hand, you may spend less time thinking and planning. Computer text editors affect structure through their limited real estate. With a paper manuscript, you can spread the pages upon the desk, couch, wall, or floor. Large sections of the text can be examined at one time, to be reorganized and structured. If you use only the computer, then the working area (or real estate) is limited to what shows on the screen. The conventional screens display about twenty-four lines of text. Even the largest screens now available can display no more than about two full printed pages of text. The result is that corrections tend to be made locally, on what is visible. Large-scale restructuring of the material is more difficult to do, and therefore seldom gets done. Sometimes the same text appears in different parts of the manuscript, without being discovered by the writer. (To the writer, everything seems familiar.)

OUTLINE PROCESSORS AND HYPERTEXT

The current fad in writing aids is the outline processor, a tool designed to encourage planning and reflection on the organization of material. The writer can compress the text into an outline or expand an outline to cover the entire manuscript. Moving a heading means moving an entire section. Outline processors attempt to overcome organizational problems by allowing collapsed views of the manuscript to be examined and manipulated. But the process seems to emphasize the organization that is visible in the outline or heading structure of the manuscript, thereby deemphasizing other aspects of the work. It is characteristic of thought processes that attention to one aspect comes at the cost of decreased attention to others. What a technology makes easy to do will get done; what it hides, or makes difficult, may very well not get done.

The next step in writing technology is already visible on the horizon: hypertext.⁹ Here we have another set of possibilities, another set of difficulties, in this case for both writer and reader. Writers frequently complain that the material they are trying to explain is complex, multi-dimensional. The ideas are all interconnected, and there is no single sequence of words to convey them properly. Moreover, readers vary enormously in skill, interest, and prior knowledge. Some need expansion of the most elementary ideas, some want more technical details.¹⁰ Some wish to focus on one set of topics, others find those uninteresting. How on earth can a single document satisfy them all, especially when that document must be in a linear sequence, words following words, chapters following chapters? It has always been considered part of the skill of a writer to be able to take otherwise chaotic material and order it appropriately for the reader. Hypertext relieves the author of this burden. In theory, it also frees the reader from the constraints of the linear order; the reader can pursue the material in whatever order seems most relevant or interesting.

Hypertext makes a virtue out of lack of organization, allowing ideas and thoughts to be juxtaposed at will. The writer throws out the ideas, attaching them to the page where they seem first relevant. The reader can take any path at all through the book. See an interesting word on the page, point at it, and the word expands into text. See a word you don't understand, and a touch gives the definition. Who could be against such a wonderful idea?

Imagine that this book was in hypertext. How would it work? Well, I've used several devices that relate to hypertext: one is the footnote,¹¹ another is parenthetical comments, and yet another is contrasting text. (I have tended not to use parenthetical asides in this book because I fear they distract, make the sentences longer, and add to the reader's memory burden, as this parenthetical statement demonstrates.)

Contrasting text, when used as a commentary, is a kind of hypertext. Here is a comment on the text itself, optional and not essential to a first reading. The typography gives signals to the reader.

Actual hypertext will be written and read using a computer, of course, so that this commentary wouldn't be visible unless it had been requested.

A footnote is essentially a signal that some comment is available to the reader. In hypertext, actual numbered footnotes will not be needed,

but some sort of signal is still required. With hypertext, the signal that more information is available can be given through color, motion (such as flashing), or typeface. Touch the special word and the material appears; you don't need a number.

So, what do you think of hypertext? Imagine trying to write something using it. The extra freedom also poses extra requirements. If hypertext really becomes available, especially in the fancy versions now being talked about—where words, sounds, video, computer graphics, simulations, and more are all available at the touch of the screen—well, it is hard to imagine anyone capable of preparing the material. It will take teams of people. I predict that there will be much experimentation, and much failure, before the dimensions of this new technology are fully explored and understood.

One thing that does bother me, however, is the belief that hypertext will save the author from having to put material in linear order. Wrong. To think this is to allow for sloppiness in writing and presentation. It is hard work to organize material, but that effort on the part of the writer is essential for the ease of the reader. Take away the need for this discipline and I fear that you pass the burden on to the reader, who may not be able to cope, and may not care to try. The advent of hypertext is apt to make writing much more difficult, not easier. Good writing, that is.

THE HOME OF THE FUTURE: A PLACE OF COMFORT OR A NEW SOURCE OF FRUSTRATION

Even as this book is being completed, new sources of pleasure and frustration are entering our lives. Two developments are worthy of mention, both intended to serve the ever-promised "house of the future." One most wonderful development is the "smart house," the place where your every want is taken care of by intelligent, omniscient appliances. The other promised development is the house of knowledge: whole libraries available at our fingertips, the world's information resources available through our telephone/television set/home computer/rooftop satellite antenna. Both developments have great potential to transform lives in just the positive ways promised, but they are also apt to explode every fear and complexity discussed in this book into reality a thousand-times over.

Imagine all of our electric appliances connected together via an intelligent "information bus." This bus (the technical term for a set of wires that acts as communication channels among devices) allows home lamps, ovens, and dishwashers to talk to one another. The central home computer senses the car pulling into the driveway, so it signals the front door to unlock, the hall lights to go on, and the oven to start preparing the meal. By the time you arrive in the house, your television set has already turned on to your favorite news station, your favorite appetizer is available in the kitchen, and the cooking of the meal has begun. Some of these systems "speak" to you (with voice-synthesizers inside their computer brains), most have sensors that detect room temperature, the outside weather, and the presence of people. All assume a master controlling device through which the house occupants inform the system of their every want. Many allow for telephone control. Going to miss your favorite show on television? Call home and instruct your VCR to record it for you. Coming home an hour later than expected? Call your home oven and delay the starting time of the meal.

Can you imagine what it would take to control these devices? How would you tell your oven when to turn on? Would you do this through the buttons available at your friendly pay telephone? Or would you lug around a portable controlling unit? In either case, the complexity boggles the mind. Do the designers of these systems have some secret cure for the problems described throughout this book or have they perhaps already mastered the lessons within? Hardly. An article entitled "The 'smartest house' in America" in the technical magazine for design engineers, *Design News*,¹² shows the normal set of arbitrary control devices, overly complex panels, and conventional computer screens and keyboards. The modern cooktop (accompanied by the caption "for the ultimate chef") has two gas burners, four electric burners, and a barbecue grill controlled through a row of eight identical-looking, evenly spaced knobs.

It is easy to imagine positive uses for intelligent home appliances. The energy-saving virtues of a home that turns on the heat only for rooms that are occupied, or waters the yard only when the ground is dry and rain does not threaten, seem virtuous indeed. Not the most critical of the problems facing humankind, perhaps, but reassuring nonetheless. But it is difficult to see how the complex instructions required for such a system will be conveyed. I find it difficult to instruct my children how to do these tasks appropriately and I often fail at them myself. How will I manage the precise, clear instructions required for

my intelligent dishwasher, especially through the very limited control mechanism I am sure to be provided with? I do not look forward to the day.

Now consider the information world of the future. The modern laser disk is capable of holding billions of characters of information.¹³ This means that instead of purchasing individual books, we can now purchase whole libraries. One compact disk can hold hundreds of thousands (even millions) of printed pages of information. Whole encyclopedias can be available at our fingertips, through our computer terminals and television screens. And when every home is connected to a central computer system through improved capacity telephone lines, or the cable television wire, or a rooftop antenna aimed at the neighborhood earth satellite, the information of the world is available to all.

There are two costs for these pleasures. One is economic: it may only cost a few dollars to manufacture a compact disk that contains the contents of one hundred books, but the cost to the consumer will be measured in the hundreds of dollars. After all, each book took an author several years of effort and a publishing house with editors and book designers another three to nine months. Connection to the world's libraries through the telephone, television, and satellite lines of the world cost money to the telephone, cable, and communication companies. These costs have to be recovered. Those of us who use the computer library search facilities available today know that it is most convenient to have them available but that each second of use is marked by the tension that the costs are piling up. Stop to reflect on something, and your bill increases astronomically. The true costs of these systems are high, and the user's continual thought that each use exacts a cost is not reassuring.

The second cost is the difficulty of finding anything in such large data bases. I can't always find my car keys or the book I was reading last night. When I read an interesting article and store it away in my files for some unknown but probable future use, I know at the time I stick it away that I may never be able to remember where I put it. If I already have these difficulties with my own limited possessions and books, imagine what it will be like when trying to find something in the libraries and data bases of the world, where the organization was done by someone else who had no idea of what my needs were. Chaos. Sheer chaos.

The society of the future: something to look forward to with pleasure, contemplation, and dread.

That design affects society is hardly news to designers. Many take the implications of their work seriously. But the conscious manipulation of society has severe drawbacks, not the least of which is the fact that not everyone agrees on the appropriate goals. Design, therefore, takes on political significance; indeed, design philosophies vary in important ways across political systems. In Western cultures, design has reflected the capitalistic importance of the marketplace, with an emphasis on exterior features deemed to be attractive to the purchaser. In the consumer economy taste is not the criterion in the marketing of expensive foods or drinks, usability is not the primary criterion in the marketing of home and office appliances. We are surrounded with objects of desire, not objects of use.¹⁴

Everyday tasks are not difficult because of their inherent complexity. They are difficult only because they require learning arbitrary relationships and arbitrary mappings, and because they sometimes require precision in their execution. The difficulties can be avoided through design that makes obvious what operations are necessary. Good design exploits constraints so that the user feels as if there is only one possible thing to do—the right thing, of course. The designer has to understand and exploit natural constraints of all kinds.

Errors are an unavoidable part of everyday life. Proper design can help decrease the incidence and severity of errors by eliminating the causes of some, minimizing the possibilities of others, and helping to make errors discoverable, once they have been made. Such design exploits the power of constraints and makes use of forcing functions and visible outcomes of actions. We do not have to experience confusion or suffer from undiscovered errors. Proper design can make a difference in our quality of life.

Now you are on your own. If you are a designer, help fight the battle for usability. If you are a user, then join your voice with those who cry for usable products. Write to manufacturers. Boycott unusable designs. Support good designs by purchasing them, even if it means going out of your way, even if it means spending a bit more. And voice your concerns to the stores that carry the products; manufacturers listen to their customers.

When you visit museums of science and technology, ask questions

if you have trouble understanding. Provide feedback about the exhibits and whether they work well or poorly. Encourage museums to move toward better usability and understandability.

And enjoy yourself. Walk around the world examining the details of design. Take pride in the little things that help; think kindly of the person who so thoughtfully put them in. Realize that even details matter, that the designer may have had to fight to include something helpful. Give mental prizes to those who practice good design: send flowers. Jeer those who don't: send weeds.