

Copyright Notice

Staff and students of UWE Bristol are reminded that copyright subsists in this extract and the work from which it was taken. This Digital Copy has been made under the terms of a CLA licence which allows you to:

- * access and download a copy
- * print out a copy

This Digital Copy and any digital or printed copy supplied to or made by you under the terms of this Licence are for use in connection with this Course of Study. You may retain such copies after the end of the course, but strictly for your own personal use.

All copies (including electronic copies) shall include this Copyright Notice and shall be destroyed and/or deleted if and when required by UWE Bristol.

Except as provided for by copyright law, no further copying, storage or distribution (including by e-mail) is permitted without the consent of the copyright holder.

The author (which term includes artists and other visual creators) has moral rights in the work and neither staff nor students may cause, or permit, the distortion, mutilation or other modification of the work, or any other derogatory treatment of it, which would be prejudicial to the honour or reputation of the author.

Course of Study:

(UFCF7L-15-1) Design thinking

Title:

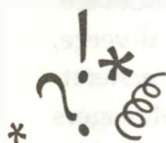
The design of everyday things pp. 217--224, 7--25

Name of Author:

Donald A. Norman

Name of Publisher:

MIT Press

DESIGN
THINKING

One of my rules in consulting is simple: never solve the problem I am asked to solve. Why such a counterintuitive rule? Because, invariably, the problem I am asked to solve is not the real, fundamental, root problem. It is usually a symptom. Just as in Chapter 5, where the solution to accidents and errors was to determine the real, underlying cause of the events, in design, the secret to success is to understand what the real problem is.

It is amazing how often people solve the problem before them without bothering to question it. In my classes of graduate students in both engineering and business, I like to give them a problem to solve on the first day of class and then listen the next week to their wonderful solutions. They have masterful analyses, drawings, and illustrations. The MBA students show spreadsheets in which they have analyzed the demographics of the potential customer base. They show lots of numbers: costs, sales, margins, and profits. The engineers show detailed drawings and specifications. It is all well done, brilliantly presented.

When all the presentations are over, I congratulate them, but ask: "How do you know you solved the correct problem?" They are puzzled. Engineers and business people are trained to solve

problems. Why would anyone ever give them the wrong problem? "Where do you think the problems come from?" I ask. The real world is not like the university. In the university, professors make up artificial problems. In the real world, the problems do not come in nice, neat packages. They have to be discovered. It is all too easy to see only the surface problems and never dig deeper to address the real issues.

Solving the Correct Problem

Engineers and businesspeople are trained to solve problems. Designers are trained to discover the real problems. A brilliant solution to the wrong problem can be worse than no solution at all: solve the correct problem.

Good designers never start by trying to solve the problem given to them: they start by trying to understand what the real issues are. As a result, rather than converge upon a solution, they diverge, studying people and what they are trying to accomplish, generating idea after idea after idea. It drives managers crazy. Managers want to see progress: designers seem to be going backward when they are given a precise problem and instead of getting to work, they ignore it and generate new issues to consider, new directions to explore. And not just one, but many. What is going on?

The key emphasis of this book is the importance of developing products that fit the needs and capabilities of people. Design can be driven by many different concerns. Sometimes it is driven by technology, sometimes by competitive pressures or by aesthetics. Some designs explore the limits of technological possibilities; some explore the range of imagination, of society, of art or fashion. Engineering design tends to emphasize reliability, cost, and efficiency. The focus of this book, and of the discipline called human-centered design, is to ensure that the result fits human desires, needs, and capabilities. After all, why do we make products? We make them for people to use.

Designers have developed a number of techniques to avoid being captured by too facile a solution. They take the original problem

as a suggestion, not as a final statement, then think broadly about what the issues underlying this problem statement might really be (as was done through the “Five Whys” approach to getting at the root cause, described in Chapter 5). Most important of all is that the process be iterative and expansive. Designers resist the temptation to jump immediately to a solution for the stated problem. Instead, they first spend time determining what basic, fundamental (root) issue needs to be addressed. They don’t try to search for a solution until they have determined the real problem, and even then, instead of solving that problem, they stop to consider a wide range of potential solutions. Only then will they finally converge upon their proposal. This process is called *design thinking*.

Design thinking is not an exclusive property of designers—all great innovators have practiced this, even if unknowingly, regardless of whether they were artists or poets, writers or scientists, engineers or businesspeople. But because designers pride themselves on their ability to innovate, to find creative solutions to fundamental problems, design thinking has become the hallmark of the modern design firm. Two of the powerful tools of design thinking are human-centered design and the double-diamond diverge-converge model of design.

Human-centered design (HCD) is the process of ensuring that people’s needs are met, that the resulting product is understandable and usable, that it accomplishes the desired tasks, and that the experience of use is positive and enjoyable. Effective design needs to satisfy a large number of constraints and concerns, including shape and form, cost and efficiency, reliability and effectiveness, understandability and usability, the pleasure of the appearance, the pride of ownership, and the joy of actual use. HCD is a procedure for addressing these requirements, but with an emphasis on two things: solving the right problem, and doing so in a way that meets human needs and capabilities.

Over time, the many different people and industries that have been involved in design have settled upon a common set of methods for doing HCD. Everyone has his or her own favorite method,

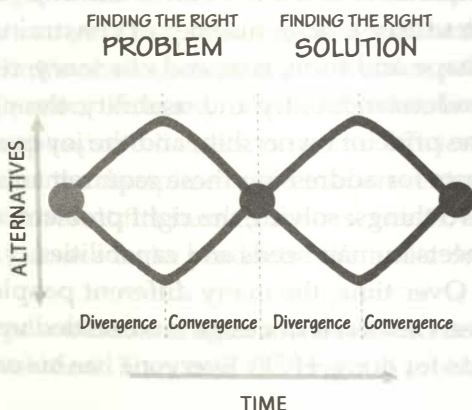
but all are variants on the common theme: iterate through the four stages of observation, generation, prototyping, and testing. But even before this, there is one overriding principle: solve the right problem.

These two components of design—finding the right problem and meeting human needs and capabilities—give rise to two phases of the design process. The first phase is to find the right problem, the second is to find the right solution. Both phases use the HCD process. This double-phase approach to design led the British Design Council to describe it as a “double diamond.” So that is where we start the story.

The Double-Diamond Model of Design

Designers often start by questioning the problem given to them: they expand the scope of the problem, diverging to examine all the fundamental issues that underlie it. Then they converge upon a single problem statement. During the solution phase of their studies, they first expand the space of possible solutions, the divergence phase. Finally, they converge upon a proposed solution (Figure 6.1). This double diverge-converge pattern was first introduced in 2005 by the British Design Council, which called it the *double-diamond design process model*. The Design Council divided the design process into four stages: “discover” and “define”—for the divergence and convergence phases of finding the right problem,

FIGURE 6.1. The Double-Diamond Model of Design. Start with an idea, and through the initial design research, expand the thinking to explore the fundamental issues. Only then is it time to converge upon the real, underlying problem. Similarly, use design research tools to explore a wide variety of solutions before converging upon one. (Slightly modified from the work of the British Design Council, 2005.)



and “develop” and “deliver”—for the divergence and convergence phases of finding the right solution.

The double diverge-converge process is quite effective at freeing designers from unnecessary restrictions to the problem and solution spaces. But you can sympathize with a product manager who, having given the designers a problem to solve, finds them questioning the assignment and insisting on traveling all over the world to seek deeper understanding. Even when the designers start focusing upon the problem, they do not seem to make progress, but instead develop a wide variety of ideas and thoughts, many only half-formed, many clearly impractical. All this can be rather unsettling to the product manager who, concerned about meeting the schedule, wants to see immediate convergence. To add to the frustration of the product manager, as the designers start to converge upon a solution, they may realize that they have inappropriately formulated the problem, so the entire process must be repeated (although it can go more quickly this time).

This repeated divergence and convergence is important in properly determining the right problem to be solved and then the best way to solve it. It looks chaotic and ill-structured, but it actually follows well-established principles and procedures. How does the product manager keep the entire team on schedule despite the apparent random and divergent methods of designers? Encourage their free exploration, but hold them to the schedule (and budget) constraints. There is nothing like a firm deadline to get creative minds to reach convergence.

The Human-Centered Design Process

The double-diamond describes the two phases of design: finding the right problem and fulfilling human needs. But how are these actually done? This is where the human-centered design process comes into play: it takes place within the double-diamond diverge-converge process.

There are four different activities in the human-centered design process (Figure 6.2):

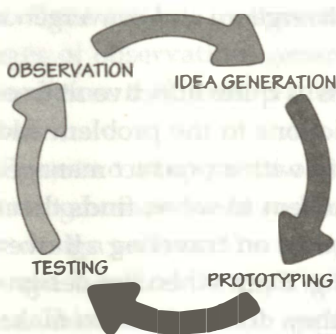


FIGURE 6.2. The Iterative Cycle of Human-Centered Design. Make observations on the intended target population, generate ideas, produce prototypes and test them. Repeat until satisfied. This is often called the *spiral method* (rather than the circle depicted here), to emphasize that each iteration through the stages makes progress.

1. Observation
2. Idea generation (ideation)
3. Prototyping
4. Testing

These four activities are iterated; that is, they are repeated over and over, with each cycle yielding more insights and getting closer to the desired solution. Now let us examine each activity separately.

OBSERVATION

The initial research to understand the nature of the problem itself is part of the discipline of design research. Note that this is research about the customer and the people who will use the products under consideration. It is not the kind of research that scientists do in their laboratories, trying to find new laws of nature. The design researcher will go to the potential customers, observing their activities, attempting to understand their interests, motives, and true needs. The problem definition for the product design will come from this deep understanding of the goals the people are trying to accomplish and the impediments they experience. One of its most critical techniques is to observe the would-be customers in their natural environment, in their normal lives, wherever the product or service being designed will actually be used. Watch them in their homes, schools, and offices. Watch them commute, at parties, at mealtime, and with friends at the local bar. Follow them into the shower if necessary, because it is essential to understand the real situations that they encounter, not some pure isolated experience. This technique is called *applied ethnography*, a method adapted from the field of anthropology. Applied ethnography differs from the slower, more methodical, research-oriented practice of academic anthropologists because the goals are different.

For one, design researchers have the goal of determining human needs that can be addressed through new products. For another, product cycles are driven by schedule and budget, both of which require more rapid assessment than is typical in academic studies that might go on for years.

It's important that the people being observed match those of the intended audience. Note that traditional measures of people, such as age, education, and income, are not always important: what matters most are the activities to be performed. Even when we look at widely different cultures, the activities are often surprisingly similar. As a result, the studies can focus upon the activities and how they get done, while being sensitive to how the local environment and culture might modify those activities. In some cases, such as the products widely used in business, the activity dominates. Thus, automobiles, computers, and phones are pretty standardized across the world because their designs reflect the activities being supported.

In some cases, detailed analyses of the intended group are necessary. Japanese teenage girls are quite different from Japanese women, and in turn, very different from German teenage girls. If a product is intended for subcultures like these, the exact population must be studied. Another way of putting it is that different products serve different needs. Some products are also symbols of status or group membership. Here, although they perform useful functions, they are also fashion statements. This is where teenagers in one culture differ from those of another, and even from younger children and older adults of the same culture. Design researchers must carefully adjust the focus of their observations to the intended market and people for whom the product is intended.

Will the product be used in some country other than where it is being designed? There is only one way to find out: go there (and always include natives in the team). Don't take a shortcut and stay home, talking to students or visitors from that country while remaining in your own: what you will learn is seldom an accurate reflection of the target population or of the ways in which the proposed product will actually be used. There is no substitute for

direct observation of and interaction with the people who will be using the product.

Design research supports both diamonds of the design process. The first diamond, finding the right problem, requires a deep understanding of the true needs of people. Once the problem has been defined, finding an appropriate solution again requires deep understanding of the intended population, how those people perform their activities, their capabilities and prior experience, and what cultural issues might be impacted.

DESIGN RESEARCH VERSUS MARKET RESEARCH

Design and marketing are two important parts of the product development group. The two fields are complementary, but each has a different focus. Design wants to know what people really need and how they actually will use the product or service under consideration. Marketing wants to know what people will buy, which includes learning how they make their purchasing decisions. These different aims lead the two groups to develop different methods of inquiry. Designers tend to use qualitative observational methods by which they can study people in depth, understanding how they do their activities and the environmental factors that come into play. These methods are very time consuming, so designers typically only examine small numbers of people, often numbering in the tens.

Marketing is concerned with customers. Who might possibly purchase the item? What factors might entice them to consider and purchase a product? Marketing traditionally uses large-scale, quantitative studies, with heavy reliance on focus groups, surveys, and questionnaires. In marketing, it is not uncommon to converse with hundreds of people in focus groups, and to question tens of thousands of people by means of questionnaires and surveys.

The advent of the Internet and the ability to assess huge amounts of data have given rise to new methods of formal, quantitative market analysis. "Big data," it is called, or sometimes "market analytics." For popular websites, A/B testing is possible in which two potential variants of an offering are tested by giving

I used to be an engineer, focused upon technical requirements, quite ignorant of people. Even after I switched into psychology and cognitive science, I still maintained my engineering emphasis upon logic and mechanism. It took a long time for me to realize that my understanding of human behavior was relevant to my interest in the design of technology. As I watched people struggle with technology, it became clear that the difficulties were caused by the technology, not the people.

I was called upon to help analyze the American nuclear power plant accident at Three Mile Island (the island name comes from the fact that it is located on a river, three miles south of Middletown in the state of Pennsylvania). In this incident, a rather simple mechanical failure was misdiagnosed. This led to several days of difficulties and confusion, total destruction of the reactor, and a very close call to a severe radiation release, all of which brought the American nuclear power industry to a complete halt. The operators were blamed for these failures: "human error" was the immediate analysis. But the committee I was on discovered that the plant's control rooms were so poorly designed that error was inevitable: design was at fault, not the operators. The moral was simple: we were designing things for people, so we needed to understand both technology and people. But that's a difficult step for many engineers: machines are so logical, so orderly. If we didn't have people, everything would work so much better. Yup, that's how I used to think.

My work with that committee changed my view of design. Today, I realize that design presents a fascinating interplay of technology and psychology, that the designers must understand both. Engineers still tend to believe in logic. They often explain to me in great, logical detail, why their designs are good, powerful, and wonderful. "Why are people having problems?" they wonder. "You are being too logical," I say. "You are designing for people the way you would like them to be, not for the way they really are."

When the engineers object, I ask whether they have ever made an error, perhaps turning on or off the wrong light, or the wrong

stove burner. "Oh yes," they say, "but those were errors." That's the point: even experts make errors. So we must design our machines on the assumption that people will make errors. (Chapter 5 provides a detailed analysis of human error.)

Human-Centered Design

People are frustrated with everyday things. From the ever-increasing complexity of the automobile dashboard, to the increasing automation in the home with its internal networks, complex music, video, and game systems for entertainment and communication, and the increasing automation in the kitchen, everyday life sometimes seems like a never-ending fight against confusion, continued errors, frustration, and a continual cycle of updating and maintaining our belongings.

In the multiple decades that have elapsed since the first edition of this book was published, design has gotten better. There are now many books and courses on the topic. But even though much has improved, the rapid rate of technology change outpaces the advances in design. New technologies, new applications, and new methods of interaction are continually arising and evolving. New industries spring up. Each new development seems to repeat the mistakes of the earlier ones; each new field requires time before it, too, adopts the principles of good design. And each new invention of technology or interaction technique requires experimentation and study before the principles of good design can be fully integrated into practice. So, yes, things are getting better, but as a result, the challenges are ever present.

The solution is human-centered design (HCD), an approach that puts human needs, capabilities, and behavior first, then designs to accommodate those needs, capabilities, and ways of behaving. Good design starts with an understanding of psychology and technology. Good design requires good communication, especially from machine to person, indicating what actions are possible, what is happening, and what is about to happen. Communication is especially important when things go wrong. It is relatively easy to design things that work smoothly and harmoniously as

TABLE 1.1. The Role of HCD and Design Specializations	
Experience design	These are areas of focus
Industrial design	
Interaction design	
Human-centered design	The process that ensures that the designs match the needs and capabilities of the people for whom they are intended

long as things go right. But as soon as there is a problem or a misunderstanding, the problems arise. This is where good design is essential. Designers need to focus their attention on the cases where things go wrong, not just on when things work as planned. Actually, this is where the most satisfaction can arise: when something goes wrong but the machine highlights the problems, then the person understands the issue, takes the proper actions, and the problem is solved. When this happens smoothly, the collaboration of person and device feels wonderful.

Human-centered design is a design philosophy. It means starting with a good understanding of people and the needs that the design is intended to meet. This understanding comes about primarily through observation, for people themselves are often unaware of their true needs, even unaware of the difficulties they are encountering. Getting the specification of the thing to be defined is one of the most difficult parts of the design, so much so that the HCD principle is to avoid specifying the problem as long as possible but instead to iterate upon repeated approximations. This is done through rapid tests of ideas, and after each test modifying the approach and the problem definition. The results can be products that truly meet the needs of people. Doing HCD within the rigid time, budget, and other constraints of industry can be a challenge: Chapter 6 examines these issues.

Where does HCD fit into the earlier discussion of the several different forms of design, especially the areas called industrial, interaction, and experience design? These are all compatible. HCD is a philosophy and a set of procedures, whereas the others are areas of focus (see Table 1.1). The philosophy and procedures of HCD add

deep consideration and study of human needs to the design process, whatever the product or service, whatever the major focus.

Fundamental Principles of Interaction

Great designers produce pleasurable experiences. *Experience*: note the word. Engineers tend not to like it; it is too subjective. But when I ask them about their favorite automobile or test equipment, they will smile delightedly as they discuss the fit and finish, the sensation of power during acceleration, their ease of control while shifting or steering, or the wonderful feel of the knobs and switches on the instrument. Those are experiences.

Experience is critical, for it determines how fondly people remember their interactions. Was the overall experience positive, or was it frustrating and confusing? When our home technology behaves in an uninterpretable fashion we can become confused, frustrated, and even angry—all strong negative emotions. When there is understanding it can lead to a feeling of control, of mastery, and of satisfaction or even pride—all strong positive emotions. Cognition and emotion are tightly intertwined, which means that the designers must design with both in mind.

When we interact with a product, we need to figure out how to work it. This means discovering what it does, how it works, and what operations are possible: discoverability. Discoverability results from appropriate application of five fundamental psychological concepts covered in the next few chapters: *affordances*, *signifiers*, *constraints*, *mappings*, and *feedback*. But there is a sixth principle, perhaps most important of all: the *conceptual model* of the system. It is the conceptual model that provides true understanding. So I now turn to these fundamental principles, starting with affordances, signifiers, mappings, and feedback, then moving to conceptual models. Constraints are covered in Chapters 3 and 4.

AFFORDANCES

We live in a world filled with objects, many natural, the rest artificial. Every day we encounter thousands of objects, many of them new to us. Many of the new objects are similar to ones we already

know, but many are unique, yet we manage quite well. How do we do this? Why is it that when we encounter many unusual natural objects, we know how to interact with them? Why is this true with many of the artificial, human-made objects we encounter? The answer lies with a few basic principles. Some of the most important of these principles come from a consideration of affordances.

The term *affordance* refers to the relationship between a physical object and a person (or for that matter, any interacting agent, whether animal or human, or even machines and robots). An affordance is a relationship between the properties of an object and the capabilities of the agent that determine just how the object could possibly be used. A chair affords (“is for”) support and, therefore, affords sitting. Most chairs can also be carried by a single person (they afford lifting), but some can only be lifted by a strong person or by a team of people. If young or relatively weak people cannot lift a chair, then for these people, the chair does not have that affordance, it does not afford lifting.

The presence of an affordance is jointly determined by the qualities of the object and the abilities of the agent that is interacting. This relational definition of affordance gives considerable difficulty to many people. We are used to thinking that properties are associated with objects. But affordance is not a property. An affordance is a relationship. Whether an affordance exists depends upon the properties of both the object and the agent.

Glass affords transparency. At the same time, its physical structure blocks the passage of most physical objects. As a result, glass affords seeing through and support, but not the passage of air or most physical objects (atomic particles can pass through glass). The blockage of passage can be considered an anti-affordance—the prevention of interaction. To be effective, affordances and anti-affordances have to be discoverable—perceivable. This poses a difficulty with glass. The reason we like glass is its relative invisibility, but this aspect, so useful in the normal window, also hides its anti-affordance property of blocking passage. As a result, birds often try to fly through windows. And every year, numerous people injure themselves when they walk (or run) through closed glass

doors or large picture windows. If an affordance or anti-affordance cannot be perceived, some means of signaling its presence is required: I call this property a *signifier* (discussed in the next section).

The notion of affordance and the insights it provides originated with J. J. Gibson, an eminent psychologist who provided many advances to our understanding of human perception. I had interacted with him over many years, sometimes in formal conferences and seminars, but most fruitfully over many bottles of beer, late at night, just talking. We disagreed about almost everything. I was an engineer who became a cognitive psychologist, trying to understand how the mind works. He started off as a Gestalt psychologist, but then developed an approach that is today named after him: Gibsonian psychology, an ecological approach to perception. He argued that the world contained the clues and that people simply picked them up through “direct perception.” I argued that nothing could be direct: the brain had to process the information arriving at the sense organs to put together a coherent interpretation. “Nonsense,” he loudly proclaimed; “it requires no interpretation: it is directly perceived.” And then he would put his hand to his ears, and with a triumphant flourish, turn off his hearing aids: my counterarguments would fall upon deaf ears—literally.

When I pondered my question—how do people know how to act when confronted with a novel situation—I realized that a large part of the answer lay in Gibson’s work. He pointed out that all the senses work together, that we pick up information about the world by the combined result of all of them. “Information pickup” was one of his favorite phrases, and Gibson believed that the combined information picked up by all of our sensory apparatus—sight, sound, smell, touch, balance, kinesthetic, acceleration, body position—determines our perceptions without the need for internal processing or cognition. Although he and I disagreed about the role played by the brain’s internal processing, his brilliance was in focusing attention on the rich amount of information present in the world. Moreover, the physical objects conveyed important information about how people could interact with them, a property he named “affordance.”

Affordances exist even if they are not visible. For designers, their visibility is critical: visible affordances provide strong clues to the operations of things. A flat plate mounted on a door affords pushing. Knobs afford turning, pushing, and pulling. Slots are for inserting things into. Balls are for throwing or bouncing. Perceived affordances help people figure out what actions are possible without the need for labels or instructions. I call the signaling component of affordances *signifiers*.

SIGNIFIERS

Are affordances important to designers? The first edition of this book introduced the term *affordances* to the world of design. The design community loved the concept and affordances soon propagated into the instruction and writing about design. I soon found mention of the term everywhere. Alas, the term became used in ways that had nothing to do with the original.

Many people find affordances difficult to understand because they are relationships, not properties. Designers deal with fixed properties, so there is a temptation to say that the property is an affordance. But that is not the only problem with the concept of affordances.

Designers have practical problems. They need to know how to design things to make them understandable. They soon discovered that when working with the graphical designs for electronic displays, they needed a way to designate which parts could be touched, slid upward, downward, or sideways, or tapped upon. The actions could be done with a mouse, stylus, or fingers. Some systems responded to body motions, gestures, and spoken words, with no touching of any physical device. How could designers describe what they were doing? There was no word that fit, so they took the closest existing word—*affordance*. Soon designers were saying such things as, “I put an affordance there,” to describe why they displayed a circle on a screen to indicate where the person should touch, whether by mouse or by finger. “No,” I said, “that is not an affordance. That is a way of communicating where the touch should be. You are communicating where to do the touching: the

affordance of touching exists on the entire screen: you are trying to signify *where* the touch should take place. That's not the same thing as saying *what* action is possible."

Not only did my explanation fail to satisfy the design community, but I myself was unhappy. Eventually I gave up: designers needed a word to describe what they were doing, so they chose *affordance*. What alternative did they have? I decided to provide a better answer: *signifiers*. Affordances determine what actions are possible. Signifiers communicate where the action should take place. We need both.

People need some way of understanding the product or service they wish to use, some sign of what it is for, what is happening, and what the alternative actions are. People search for clues, for any sign that might help them cope and understand. It is the sign that is important, anything that might signify meaningful information. Designers need to provide these clues. What people need, and what designers must provide, are signifiers. Good design requires, among other things, good communication of the purpose, structure, and operation of the device to the people who use it. That is the role of the signifier.

The term *signifier* has had a long and illustrious career in the exotic field of semiotics, the study of signs and symbols. But just as I appropriated *affordance* to use in design in a manner somewhat different than its inventor had intended, I use *signifier* in a somewhat different way than it is used in semiotics. For me, the term *signifier* refers to any mark or sound, any perceivable indicator that communicates appropriate behavior to a person.

Signifiers can be deliberate and intentional, such as the sign *PUSH* on a door, but they may also be accidental and unintentional, such as our use of the visible trail made by previous people walking through a field or over a snow-covered terrain to determine the best path. Or how we might use the presence or absence of people waiting at a train station to determine whether we have missed the train. (I explain these ideas in more detail in my book *Living with Complexity*.)

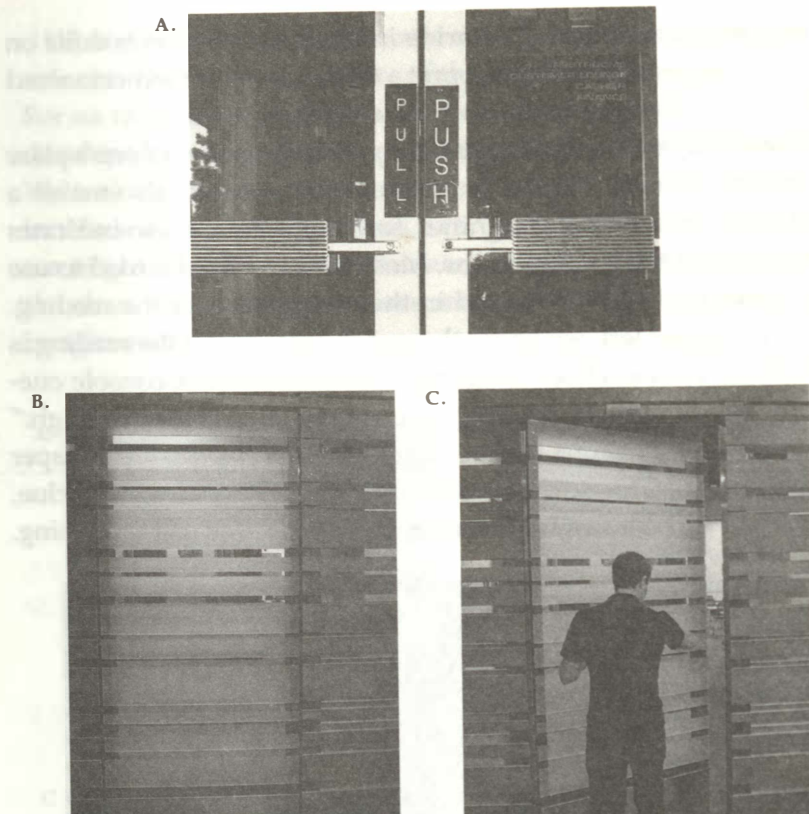


FIGURE 1.2. Problem Doors: Signifiers Are Needed. Door hardware can signal whether to push or pull without signs, but the hardware of the two doors in the upper photo, A, are identical even though one should be pushed, the other pulled. The flat, ribbed horizontal bar has the obvious perceived affordance of pushing, but as the signs indicate, the door on the left is to be pulled, the one on the right is to be pushed. In the bottom pair of photos, B and C, there are no visible signifiers or affordances. How does one know which side to push? Trial and error. When external signifiers—signs—have to be added to something as simple as a door, it indicates bad design. (Photographs by the author.)

The signifier is an important communication device to the recipient, whether or not communication was intended. It doesn't matter whether the useful signal was deliberately placed or whether it is incidental: there is no necessary distinction. Why should it matter whether a flag was placed as a deliberate clue to wind direction (as is done at airports or on the masts of sailboats) or was there as an

advertisement or symbol of pride in one's country (as is done on public buildings). Once I interpret a flag's motion to indicate wind direction, it does not matter why it was placed there.

Consider a bookmark, a deliberately placed signifier of one's place in reading a book. But the physical nature of books also makes a bookmark an accidental signifier, for its placement also indicates how much of the book remains. Most readers have learned to use this accidental signifier to aid in their enjoyment of the reading. With few pages left, we know the end is near. And if the reading is torturous, as in a school assignment, one can always console oneself by knowing there are "only a few more pages to get through." Electronic book readers do not have the physical structure of paper books, so unless the software designer deliberately provides a clue, they do not convey any signal about the amount of text remaining.

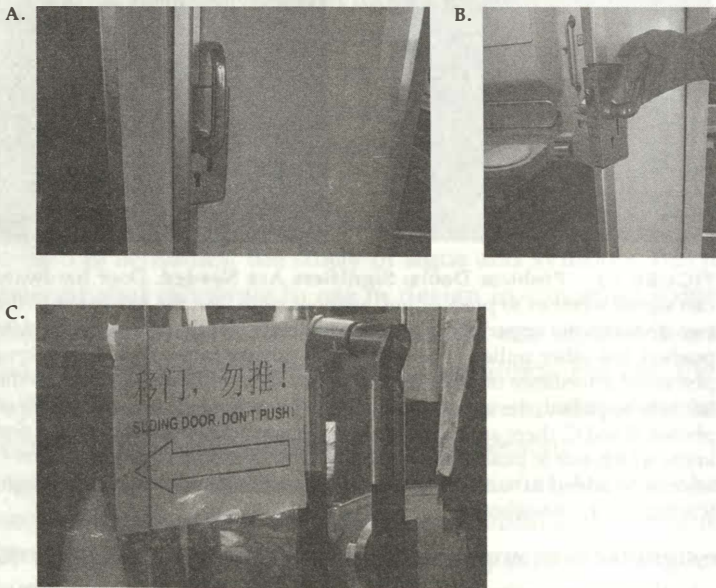


FIGURE 1.3. Sliding Doors: Seldom Done Well. Sliding doors are seldom signified properly. The top two photographs show the sliding door to the toilet on an Amtrak train in the United States. The handle clearly signifies "pull," but in fact, it needs to be rotated and the door slid to the right. The owner of the store in Shanghai, China, Photo C, solved the problem with a sign. "DON'T PUSH!" it says, in both English and Chinese. Amtrak's toilet door could have used a similar kind of sign. (Photographs by the author.)

Whatever their nature, planned or accidental, signifiers provide valuable clues as to the nature of the world and of social activities. For us to function in this social, technological world, we need to develop internal models of what things mean, of how they operate. We seek all the clues we can find to help in this enterprise, and in this way, we are detectives, searching for whatever guidance we might find. If we are fortunate, thoughtful designers provide the clues for us. Otherwise, we must use our own creativity and imagination.

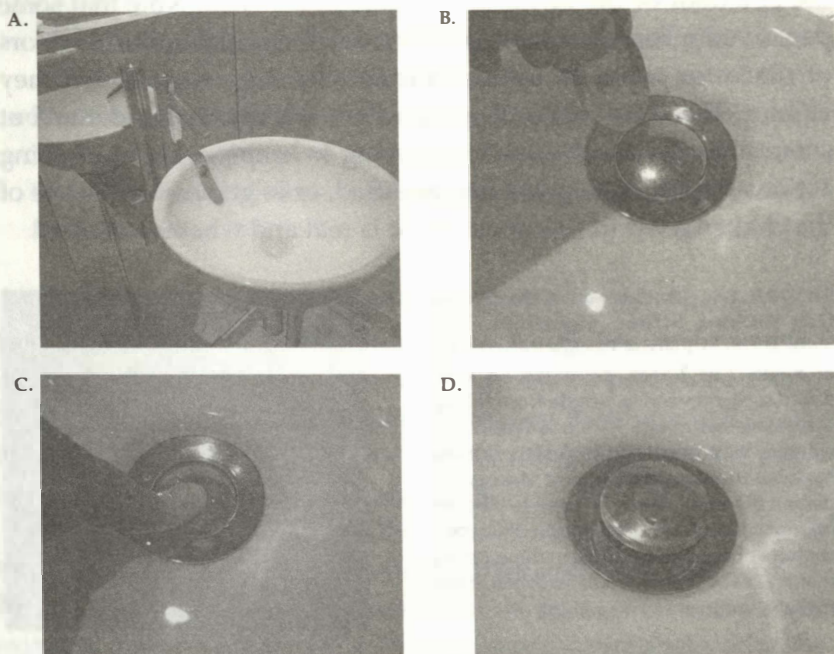


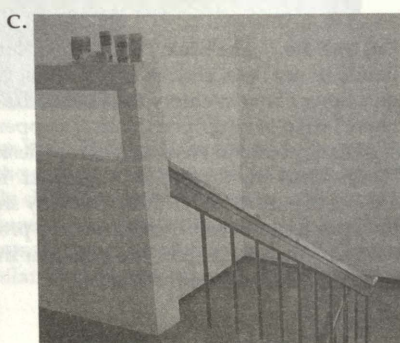
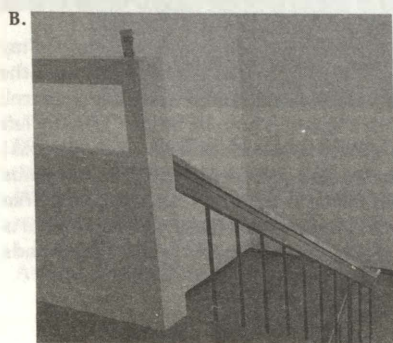
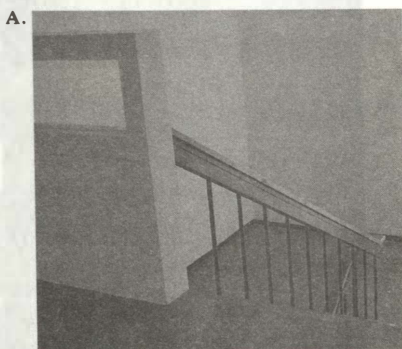
FIGURE 1.4. The Sink That Would Not Drain: Where Signifiers Fail. I washed my hands in my hotel sink in London, but then, as shown in Photo A, was left with the question of how to empty the sink of the dirty water. I searched all over for a control: none. I tried prying open the sink stopper with a spoon (Photo B): failure. I finally left my hotel room and went to the front desk to ask for instructions. (Yes, I actually did.) "Push down on the stopper," I was told. Yes, it worked (Photos C and D). But how was anyone to ever discover this? And why should I have to put my clean hands back into the dirty water to empty the sink? The problem here is not just the lack of signifier, it is the faulty decision to produce a stopper that requires people to dirty their clean hands to use it. (Photographs by the author.)

Affordances, perceived affordances, and signifiers have much in common, so let me pause to ensure that the distinctions are clear.

Affordances represent the possibilities in the world for how an agent (a person, animal, or machine) can interact with something. Some affordances are perceivable, others are invisible. Signifiers are signals. Some signifiers are signs, labels, and drawings placed in the world, such as the signs labeled “push,” “pull,” or “exit” on doors, or arrows and diagrams indicating what is to be acted upon or in which direction to gesture, or other instructions. Some signifiers are simply the perceived affordances, such as the handle of a door or the physical structure of a switch. Note that some perceived affordances may not be real: they may look like doors or places to push, or an impediment to entry, when in fact they are not. These are misleading signifiers, oftentimes accidental but sometimes purposeful, as when trying to keep people from doing actions for which they are not qualified, or in games, where one of the challenges is to figure out what is real and what is not.

FIGURE 1.5. Accidental Affordances Can Become Strong Signifiers.

This wall, at the Industrial Design department of KAIST, in Korea, provides an anti-affordance, preventing people from falling down the stair shaft. Its top is flat, an accidental by-product of the design. But flat surfaces afford support, and as soon as one person discovers it can be used to dispose of empty drink containers, the discarded container becomes a signifier, telling others that it is permissible to discard their items there. (Photographs by the author.)



My favorite example of a misleading signifier is a row of vertical pipes across a service road that I once saw in a public park. The pipes obviously blocked cars and trucks from driving on that road: they were good examples of anti-affordances. But to my great surprise, I saw a park vehicle simply go through the pipes. Huh? I walked over and examined them: the pipes were made of rubber, so vehicles could simply drive right over them. A very clever signifier, signaling a blocked road (via an apparent anti-affordance) to the average person, but permitting passage for those who knew.

To summarize:

- Affordances are the possible interactions between people and the environment. Some affordances are perceivable, others are not.
- Perceived affordances often act as signifiers, but they can be ambiguous.
- Signifiers signal things, in particular what actions are possible and how they should be done. Signifiers must be perceivable, else they fail to function.

In design, signifiers are more important than affordances, for they communicate how to use the design. A signifier can be words, a graphical illustration, or just a device whose perceived affordances are unambiguous. Creative designers incorporate the signifying part of the design into a cohesive experience. For the most part, designers can focus upon signifiers.

Because affordances and signifiers are fundamentally important principles of good design, they show up frequently in the pages of this book. Whenever you see hand-lettered signs pasted on doors, switches, or products, trying to explain how to work them, what to do and what not to do, you are also looking at poor design.

AFFORDANCES AND SIGNIFIERS: A CONVERSATION

A designer approaches his mentor. He is working on a system that recommends restaurants to people, based upon their preferences and those of their friends. But in his tests, he discovered that people never used all of the features. "Why not?" he asks his mentor. (With apologies to Socrates.)

DESIGNER	MENTOR
I'm frustrated; people aren't using our application properly.	Can you tell me about it?
The screen shows the restaurant that we recommend. It matches their preferences, and their friends like it as well. If they want to see other recommendations, all they have to do is swipe left or right. To learn more about a place, just swipe up for a menu or down to see if any friends are there now. People seem to find the other recommendations, but not the menus or their friends? I don't understand.	Why do you think this might be?
I don't know. Should I add some affordances? Suppose I put an arrow on each edge and add a label saying what they do.	That is very nice. But why do you call these affordances? They could already do the actions. Weren't the affordances already there?
Yes, you have a point. But the affordances weren't visible. I made them visible.	Very true. You added a signal of what to do.
Yes, isn't that what I said?	Not quite—you called them affordances even though they afford nothing new: they signify what to do and where to do it. So call them by their right name: " <i>signifiers</i> ."
Oh, I see. But then why do designers care about affordances? Perhaps we should focus our attention on signifiers.	You speak wisely. Communication is a key to good design. And a key to communication is the signifier.
Oh. Now I understand my confusion. Yes, a signifier is what signifies. It is a sign. Now it seems perfectly obvious.	Profound ideas are always obvious once they are understood.

MAPPING

Mapping is a technical term, borrowed from mathematics, meaning the relationship between the elements of two sets of things. Suppose there are many lights in the ceiling of a classroom or auditorium and a row of light switches on the wall at the front of the

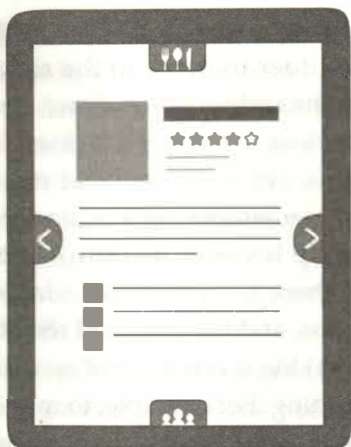


FIGURE 1.6. Signifiers on a Touch Screen. The arrows and icons are signifiers: they provide signals about the permissible operations for this restaurant guide. Swiping left or right brings up new restaurant recommendations. Swiping up reveals the menu for the restaurant being displayed; swiping down, friends who recommend the restaurant.

room. The mapping of switches to lights specifies which switch controls which light.

Mapping is an important concept in the design and layout of controls and displays. When the mapping uses spatial correspondence between the layout of the controls and the devices being controlled, it is easy to determine how to use them. In steering a car, we rotate the steering wheel clockwise to cause the car to turn right: the top of the wheel moves in the same direction as the car. Note that other choices could have been made. In early cars, steering was controlled by a variety of devices, including tillers, handlebars, and reins. Today, some vehicles use joysticks, much as in a computer game. In cars that used tillers, steering was done much as one steers a boat: move the tiller to the left to turn to the right. Tractors, construction equipment such as bulldozers and cranes, and military tanks that have tracks instead of wheels use separate controls for the speed and direction of each track: to turn right, the left track is increased in speed, while the right track is slowed or even reversed. This is also how a wheelchair is steered.

All of these mappings for the control of vehicles work because each has a compelling conceptual model of how the operation of the control affects the vehicle. Thus, if we speed up the left wheel of a wheelchair while stopping the right wheel, it is easy to imagine the chair's pivoting on the right wheel, circling to the right. In

a small boat, we can understand the tiller by realizing that pushing the tiller to the left causes the ship's rudder to move to the right and the resulting force of the water on the rudder slows down the right side of the boat, so that the boat rotates to the right. It doesn't matter whether these conceptual models are accurate: what matters is that they provide a clear way of remembering and understanding the mappings. The relationship between a control and its results is easiest to learn wherever there is an understandable mapping between the controls, the actions, and the intended result.

Natural mapping, by which I mean taking advantage of spatial analogies, leads to immediate understanding. For example, to move an object up, move the control up. To make it easy to determine which control works which light in a large room or auditorium, arrange the controls in the same pattern as the lights. Some natural mappings are cultural or biological, as in the universal standard that moving the hand up signifies more, moving it down signifies less, which is why it is appropriate to use vertical position to represent intensity or amount. Other natural mappings follow from the principles of perception and allow for the natural grouping or patterning of controls and feedback. Groupings and proximity are important principles from Gestalt psychology that can be used to map controls to function: related controls should be grouped together. Controls should be close to the item being controlled.

Note that there are many mappings that feel "natural" but in fact are specific to a particular culture: what is natural for one culture is not necessarily natural for another. In Chapter 3, I discuss how

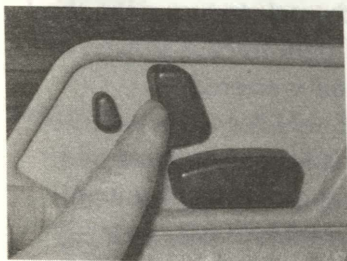


FIGURE 1.7. Good Mapping: Automobile Seat Adjustment Control. This is an excellent example of natural mapping. The control is in the shape of the seat itself: the mapping is straightforward. To move the front edge of the seat higher, lift up on the front part of the button. To make the seat back recline, move the button back. The same principle could be applied to much more common objects. This particular control is from Mercedes-Benz, but this form of mapping is now used by many automobile companies. (Photograph by the author.)

different cultures view time, which has important implications for some kinds of mappings.

A device is easy to use when the set of possible actions is visible, when the controls and displays exploit natural mappings. The principles are simple but rarely incorporated into design. Good design takes care, planning, thought, and an understanding of how people behave.

FEEDBACK

Ever watch people at an elevator repeatedly push the Up button, or repeatedly push the pedestrian button at a street crossing? Ever drive to a traffic intersection and wait an inordinate amount of time for the signals to change, wondering all the time whether the detection circuits noticed your vehicle (a common problem with bicycles)? What is missing in all these cases is feedback: some way of letting you know that the system is working on your request.

Feedback—communicating the results of an action—is a well-known concept from the science of control and information theory. Imagine trying to hit a target with a ball when you cannot see the target. Even as simple a task as picking up a glass with the hand requires feedback to aim the hand properly, to grasp the glass, and to lift it. A misplaced hand will spill the contents, too hard a grip will break the glass, and too weak a grip will allow it to fall. The human nervous system is equipped with numerous feedback mechanisms, including visual, auditory, and touch sensors, as well as vestibular and proprioceptive systems that monitor body position and muscle and limb movements. Given the importance of feedback, it is amazing how many products ignore it.

Feedback must be immediate: even a delay of a tenth of a second can be disconcerting. If the delay is too long, people often give up, going off to do other activities. This is annoying to the people, but it can also be wasteful of resources when the system spends considerable time and effort to satisfy the request, only to find that the intended recipient is no longer there. Feedback must also be informative. Many companies try to save money by using inexpensive lights or sound generators for feedback. These simple light flashes

or beeps are usually more annoying than useful. They tell us that something has happened, but convey very little information about what has happened, and then nothing about what we should do about it. When the signal is auditory, in many cases we cannot even be certain which device has created the sound. If the signal is a light, we may miss it unless our eyes are on the correct spot at the correct time. Poor feedback can be worse than no feedback at all, because it is distracting, uninformative, and in many cases irritating and anxiety-provoking.

Too much feedback can be even more annoying than too little. My dishwasher likes to beep at three a.m. to tell me that the wash is done, defeating my goal of having it work in the middle of the night so as not to disturb anyone (and to use less expensive electricity). But worst of all is inappropriate, uninterpretable feedback. The irritation caused by a "backseat driver" is well enough known that it is the staple of numerous jokes. Backseat drivers are often correct, but their remarks and comments can be so numerous and continuous that instead of helping, they become an irritating distraction. Machines that give too much feedback are like backseat drivers. Not only is it distracting to be subjected to continual flashing lights, text announcements, spoken voices, or beeps and boops, but it can be dangerous. Too many announcements cause people to ignore all of them, or wherever possible, disable all of them, which means that critical and important ones are apt to be missed. Feedback is essential, but not when it gets in the way of other things, including a calm and relaxing environment.

Poor design of feedback can be the result of decisions aimed at reducing costs, even if they make life more difficult for people. Rather than use multiple signal lights, informative displays, or rich, musical sounds with varying patterns, the focus upon cost reduction forces the design to use a single light or sound to convey multiple types of information. If the choice is to use a light, then one flash might mean one thing; two rapid flashes, something else. A long flash might signal yet another state; and a long flash followed by a brief one, yet another. If the choice is to use a sound, quite often the least expensive sound device is selected, one that

can only produce a high-frequency beep. Just as with the lights, the only way to signal different states of the machine is by beeping different patterns. What do all these different patterns mean? How can we possibly learn and remember them? It doesn't help that every different machine uses a different pattern of lights or beeps, sometimes with the same patterns meaning contradictory things for different machines. All the beeps sound alike, so it often isn't even possible to know which machine is talking to us.

Feedback has to be planned. All actions need to be confirmed, but in a manner that is unobtrusive. Feedback must also be prioritized, so that unimportant information is presented in an unobtrusive fashion, but important signals are presented in a way that does capture attention. When there are major emergencies, then even important signals have to be prioritized. When every device is signaling a major emergency, nothing is gained by the resulting cacophony. The continual beeps and alarms of equipment can be dangerous. In many emergencies, workers have to spend valuable time turning off all the alarms because the sounds interfere with the concentration required to solve the problem. Hospital operating rooms, emergency wards. Nuclear power control plants. Airplane cockpits. All can become confusing, irritating, and life-endangering places because of excessive feedback, excessive alarms, and incompatible message coding. Feedback is essential, but it has to be done correctly. Appropriately.

CONCEPTUAL MODELS

A conceptual model is an explanation, usually highly simplified, of how something works. It doesn't have to be complete or even accurate as long as it is useful. The files, folders, and icons you see displayed on a computer screen help people create the conceptual model of documents and folders inside the computer, or of apps or applications residing on the screen, waiting to be summoned. In fact, there are no folders inside the computer—those are effective conceptualizations designed to make them easier to use. Sometimes these depictions can add to the confusion, however. When reading e-mail or visiting a website, the material appears to be on