

REVISED & EXPANDED EDITION

The DESIGN
of EVERYDAY
THINGS

DON
NORMAN



THE
DESIGN
OF EVERYDAY
THINGS

ALSO BY
DON NORMAN

TEXTBOOKS

Memory and Attention: An Introduction to Human Information Processing.

First edition, 1969; second edition 1976

Human Information Processing.

(with Peter Lindsay: first edition, 1972; second edition 1977)

SCIENTIFIC MONOGRAPHS

Models of Human Memory

(edited, 1970)

Explorations in Cognition

(with David E. Rumelhart and the LNR Research Group, 1975)

Perspectives on Cognitive Science

(edited, 1981)

User Centered System Design: New Perspectives on Human-Computer Interaction

(edited with Steve Draper, 1986)

TRADE BOOKS

Learning and Memory, 1982

The Psychology of Everyday Things, 1988

The Design of Everyday Things

1990 and 2002 (paperbacks of *The Psychology of Everyday Things*
with new prefaces)

The Design of Everyday Things

Revised and Expanded Edition, 2013

Turn Signals Are the Facial Expressions of Automobiles, 1992

Things That Make Us Smart, 1993

The Invisible Computer: Why Good Products Can Fail, the Personal Computer Is So Complex, and Information Appliances Are the Answer, 1998

Emotional Design: Why We Love (or Hate) Everyday Things, 2004

The Design of Future Things, 2007

A Comprehensive Strategy for Better Reading: Cognition and Emotion, 2010

(with Masanori Okimoto; my essays, with commentary in Japanese, used
for teaching English as a second language to Japanese speakers)

Living with Complexity, 2011

CD-ROM

First person: Donald A. Norman. Defending Human Attributes in the Age of the Machine, 1994

THE DESIGN OF EVERYDAY THINGS

REVISED AND EXPANDED EDITION

Don Norman

BASIC BOOKS

A Member of the Perseus Books Group
New York

Copyright © 2013 by Don Norman

Published by Basic Books,
A Member of the Perseus Books Group

All rights reserved. Printed in the United States of America. No part of this book may be reproduced in any manner whatsoever without written permission except in the case of brief quotations embodied in critical articles and reviews. For information, address Basic Books, 250 West 57th Street, 15th Floor, New York, New York 10107.

Books published by Basic Books are available at special discounts for bulk purchases in the United States by corporations, institutions, and other organizations. For more information, please contact the Special Markets Department at the Perseus Books Group, 2300 Chestnut Street, Suite 200, Philadelphia, PA 19103, or call (800) 810-4145, ext. 5000, or e-mail special.markets@perseusbooks.com.

Library of Congress Cataloging-in-Publication Data

Norman, Donald A.

[Psychology of everyday things]

The design of everyday things / Don Norman.—Revised
and expanded edition.

pages cm

ISBN 978-0-465-05065-9 (pbk.)—ISBN 978-0-465-00394-5

(ebook) 1. Industrial design—Psychological aspects. 2. Human
engineering. I. Title.

TS171.4.N67 2013

745.2001'9—dc23

2013024417

10 9 8 7 6 5 4 3 2 1

For Julie

CONTENTS

	Preface to the Revised Edition	xi
1	The Psychopathology of Everyday Things	1
	The Complexity of Modern Devices, 4	
	Human-Centered Design, 8	
	Fundamental Principles of Interaction, 10	
	The System Image, 31	
	The Paradox of Technology, 32	
	The Design Challenge, 34	
2	The Psychology of Everyday Actions	37
	How People Do Things: The Gulfs of Execution and Evaluation, 38	
	The Seven Stages of Action, 40	
	Human Thought: Mostly Subconscious, 44	
	Human Cognition and Emotion, 49	
	The Seven Stages of Action and the Three Levels of Processing, 55	
	People as Storytellers, 56	
	Blaming the Wrong Things, 59	
	Falsely Blaming Yourself, 65	
	The Seven Stages of Action: Seven Fundamental Design Principles, 71	

3	Knowledge in the Head and in the World	74
	Precise Behavior from Imprecise Knowledge, 75	
	Memory Is Knowledge in the Head, 86	
	The Structure of Memory, 91	
	Approximate Models: Memory in the Real World, 100	
	Knowledge in the Head, 105	
	The Tradeoff Between Knowledge in the World and in the Head, 109	
	Memory in Multiple Heads, Multiple Devices, 111	
	Natural Mapping, 113	
	Culture and Design: Natural Mappings Can Vary with Culture, 118	
4	Knowing What to Do: Constraints, Discoverability, and Feedback	123
	Four Kinds of Constraints: Physical, Cultural, Semantic, and Logical, 125	
	Applying Affordances, Signifiers, and Constraints to Everyday Objects, 132	
	Constraints That Force the Desired Behavior, 141	
	Conventions, Constraints, and Affordances, 145	
	The Faucet: A Case History of Design, 150	
	Using Sound as Signifiers, 155	
5	Human Error? No, Bad Design	162
	Understanding Why There Is Error, 163	
	Deliberate Violations, 169	
	Two Types of Errors: Slips and Mistakes, 170	
	The Classification of Slips, 173	
	The Classification of Mistakes, 179	
	Social and Institutional Pressures, 186	
	Reporting Error, 191	
	Detecting Error, 194	
	Designing for Error, 198	
	When Good Design Isn't Enough, 210	
	Resilience Engineering, 211	
	The Paradox of Automation, 213	
	Design Principles for Dealing with Error, 215	

6	Design Thinking	217
	Solving the Correct Problem, 218	
	The Double-Diamond Model of Design, 220	
	The Human-Centered Design Process, 221	
	What I Just Told You? It Doesn't Really Work That Way, 236	
	The Design Challenge, 239	
	Complexity Is Good; It Is Confusion That Is Bad, 247	
	Standardization and Technology, 248	
	Deliberately Making Things Difficult, 255	
	Design: Developing Technology for People, 257	
7	Design in the World of Business	258
	Competitive Forces, 259	
	New Technologies Force Change, 264	
	How Long Does It Take to Introduce a New Product?, 268	
	Two Forms of Innovation: Incremental and Radical, 279	
	The Design of Everyday Things: 1988–2038, 282	
	The Future of Books, 288	
	The Moral Obligations of Design, 291	
	Design Thinking and Thinking About Design, 293	
	Acknowledgments	299
	General Readings and Notes	305
	References	321
	Index	331

PREFACE TO THE REVISED EDITION

In the first edition of this book, then called POET, *The Psychology of Everyday Things*, I started with these lines: “This is the book I always wanted to write, except I didn’t know it.” Today I do know it, so I simply say, “This is the book I always wanted to write.”

This is a starter kit for good design. It is intended to be enjoyable and informative for everyone: everyday people, technical people, designers, and nondesigners. One goal is to turn readers into great observers of the absurd, of the poor design that gives rise to so many of the problems of modern life, especially of modern technology. It will also turn them into observers of the good, of the ways in which thoughtful designers have worked to make our lives easier and smoother. Good design is actually a lot harder to notice than poor design, in part because good designs fit our needs so well that the design is invisible, serving us without drawing attention to itself. Bad design, on the other hand, screams out its inadequacies, making itself very noticeable.

Along the way I lay out the fundamental principles required to eliminate problems, to turn our everyday stuff into enjoyable products that provide pleasure and satisfaction. The combination of good observation skills and good design principles is a powerful

tool, one that everyone can use, even people who are not professional designers. Why? Because we are all designers in the sense that all of us deliberately design our lives, our rooms, and the way we do things. We can also design workarounds, ways of overcoming the flaws of existing devices. So, one purpose of this book is to give back your control over the products in your life: to know how to select usable and understandable ones, to know how to fix those that aren't so usable or understandable.

The first edition of the book has lived a long and healthy life. Its name was quickly changed to *Design of Everyday Things* (DOET) to make the title less cute and more descriptive. DOET has been read by the general public and by designers. It has been assigned in courses and handed out as required readings in many companies. Now, more than twenty years after its release, the book is still popular. I am delighted by the response and by the number of people who correspond with me about it, who send me further examples of thoughtless, inane design, plus occasional examples of superb design. Many readers have told me that it has changed their lives, making them more sensitive to the problems of life and to the needs of people. Some changed their careers and became designers because of the book. The response has been amazing.

Why a Revised Edition?

In the twenty-five years that have passed since the first edition of the book, technology has undergone massive change. Neither cell phones nor the Internet were in widespread usage when I wrote the book. Home networks were unheard of. Moore's law proclaims that the power of computer processors doubles roughly every two years. This means that today's computers are five thousand times more powerful than the ones available when the book was first written.

Although the fundamental design principles of *The Design of Everyday Things* are still as true and as important as when the first edition was written, the examples were badly out of date. "What is a slide projector?" students ask. Even if nothing else was to be changed, the examples had to be updated.

The principles of effective design also had to be brought up to date. Human-centered design (HCD) has emerged since the first edition, partially inspired by that book. This current edition has an entire chapter devoted to the HCD process of product development. The first edition of the book focused upon making products understandable and usable. The total experience of a product covers much more than its usability: aesthetics, pleasure, and fun play critically important roles. There was no discussion of pleasure, enjoyment, or emotion. Emotion is so important that I wrote an entire book, *Emotional Design*, about the role it plays in design. These issues are also now included in this edition.

My experiences in industry have taught me about the complexities of the real world, how cost and schedules are critical, the need to pay attention to competition, and the importance of multidisciplinary teams. I learned that the successful product has to appeal to customers, and the criteria they use to determine what to purchase may have surprisingly little overlap with the aspects that are important during usage. The best products do not always succeed. Brilliant new technologies might take decades to become accepted. To understand products, it is not enough to understand design or technology: it is critical to understand business.

What Has Changed?

For readers familiar with the earlier edition of this book, here is a brief review of the changes.

What has changed? Not much. Everything.

When I started, I assumed that the basic principles were still true, so all I needed to do was update the examples. But in the end, I rewrote everything. Why? Because although all the principles still applied, in the twenty-five years since the first edition, much has been learned. I also now know which parts were difficult and therefore need better explanations. In the interim, I also wrote many articles and six books on related topics, some of which I thought important to include in the revision. For example, the original book says nothing of what has come to be called *user experience* (a term that I was among the first to use, when in the

early 1990s, the group I headed at Apple called itself “the User Experience Architect’s Office”). This needed to be here.

Finally, my exposure to industry taught me much about the way products actually get deployed, so I added considerable information about the impact of budgets, schedules, and competitive pressures. When I wrote the original book, I was an academic researcher. Today, I have been an industry executive (Apple, HP, and some startups), a consultant to numerous companies, and a board member of companies. I had to include my learnings from these experiences.

Finally, one important component of the original edition was its brevity. The book could be read quickly as a basic, general introduction. I kept that feature unchanged. I tried to delete as much as I added to keep the total size about the same (I failed). The book is meant to be an introduction: advanced discussions of the topics, as well as a large number of important but more advanced topics, have been left out to maintain the compactness. The previous edition lasted from 1988 to 2013. If the new edition is to last as long, 2013 to 2038, I had to be careful to choose examples that would not be dated twenty-five years from now. As a result, I have tried not to give specific company examples. After all, who remembers the companies of twenty-five years ago? Who can predict what new companies will arise, what existing companies will disappear, and what new technologies will arise in the next twenty-five years? The one thing I can predict with certainty is that the principles of human psychology will remain the same, which means that the design principles here, based on psychology, on the nature of human cognition, emotion, action, and interaction with the world, will remain unchanged.

Here is a brief summary of the changes, chapter by chapter.

Chapter 1: The Psychopathology of Everyday Things

Signifiers are the most important addition to the chapter, a concept first introduced in my book *Living with Complexity*. The first edition had a focus upon affordances, but although affordances

make sense for interaction with physical objects, they are confusing when dealing with virtual ones. As a result, affordances have created much confusion in the world of design. Affordances define what actions are possible. Signifiers specify how people discover those possibilities: signifiers are signs, perceptible signals of what can be done. Signifiers are of far more importance to designers than are affordances. Hence, the extended treatment.

I added a very brief section on HCD, a term that didn't yet exist when the first edition was published, although looking back, we see that the entire book was about HCD.

Other than that, the chapter is the same, and although all the photographs and drawings are new, the examples are pretty much the same.

Chapter 2: The Psychology of Everyday Actions

The chapter has one major addition to the coverage in the first edition: the addition of emotion. The seven-stage model of action has proven to be influential, as has the three-level model of processing (introduced in my book *Emotional Design*). In this chapter I show the interplay between these two, show that different emotions arise at the different stages, and show which stages are primarily located at each of the three levels of processing (visceral, for the elementary levels of motor action performance and perception; behavioral, for the levels of action specification and initial interpretation of the outcome; and reflective, for the development of goals, plans, and the final stage of evaluation of the outcome).

Chapter 3: Knowledge in the Head and in the World

Aside from improved and updated examples, the most important addition to this chapter is a section on culture, which is of special importance to my discussion of "natural mappings." What seems natural in one culture may not be in another. The section examines the way different cultures view time—the discussion might surprise you.

Chapter. 4: Knowing What to Do: Constraints, Discoverability, and Feedback

Few substantive changes. Better examples. The elaboration of forcing functions into two kinds: lock-in and lockout. And a section on destination control elevators, illustrating how change can be extremely disconcerting, even to professionals, even if the change is for the better.

Chapter 5: Human Error? No, Bad Design

The basics are unchanged, but the chapter itself has been heavily revised. I update the classification of errors to fit advances since the publication of the first edition. In particular, I now divide slips into two main categories—action-based and memory lapses; and mistakes into three categories—rule-based, knowledge-based, and memory lapses. (These distinctions are now common, but I introduce a slightly different way to treat memory lapses.)

Although the multiple classifications of slips provided in the first edition are still valid, many have little or no implications for design, so they have been eliminated from the revision. I provide more design-relevant examples. I show the relationship of the classification of errors, slips, and mistakes to the seven-stage model of action, something new in this revision.

The chapter concludes with a quick discussion of the difficulties posed by automation (from my book *The Design of Future Things*) and what I consider the best new approach to deal with design so as to either eliminate or minimize human error: resilience engineering.

Chapter 6: Design Thinking

This chapter is completely new. I discuss two views of human-centered design: the British Design Council's double-diamond model and the traditional HCD iteration of observation, ideation, prototyping, and testing. The first diamond is the divergence, followed by convergence, of possibilities to determine the appropriate problem. The second diamond is a divergence-convergence to determine an appropriate solution. I introduce

activity-centered design as a more appropriate variant of human-centered design in many circumstances. These sections cover the theory.

The chapter then takes a radical shift in position, starting with a section entitled “What I Just Told You? It Doesn’t Really Work That Way.” Here is where I introduce Norman’s Law: The day the product team is announced, it is behind schedule and over its budget.

I discuss challenges of design within a company, where schedules, budgets, and the competing requirements of the different divisions all provide severe constraints upon what can be accomplished. Readers from industry have told me that they welcome these sections, which capture the real pressures upon them.

The chapter concludes with a discussion of the role of standards (modified from a similar discussion in the earlier edition), plus some more general design guidelines.

Chapter 7: Design in the World of Business

This chapter is also completely new, continuing the theme started in Chapter 6 of design in the real world. Here I discuss “featuritis,” the changes being forced upon us through the invention of new technologies, and the distinction between incremental and radical innovation. Everyone wants radical innovation, but the truth is, most radical innovations fail, and even when they do succeed, it can take multiple decades before they are accepted. Radical innovation, therefore, is relatively rare: incremental innovation is common.

The techniques of human-centered design are appropriate to incremental innovation: they cannot lead to radical innovations.

The chapter concludes with discussions of the trends to come, the future of books, the moral obligations of design, and the rise of small, do-it-yourself makers that are starting to revolutionize the way ideas are conceived and introduced into the marketplace: “the rise of the small,” I call it.

Summary

With the passage of time, the psychology of people stays the same, but the tools and objects in the world change. Cultures change.

Technologies change. The principles of design still hold, but the way they get applied needs to be modified to account for new activities, new technologies, new methods of communication and interaction. *The Psychology of Everyday Things* was appropriate for the twentieth century: *The Design of Everyday Things* is for the twenty-first.

Don Norman
Silicon Valley, California
www.jnd.org

THE PSYCHOPATHOLOGY OF EVERYDAY THINGS



If I were placed in the cockpit of a modern jet airliner, my inability to perform well would neither surprise nor bother me. But why should I have trouble with doors and light switches, water faucets and stoves? “Doors?” I can hear the reader saying. “You have trouble opening doors?” Yes. I push doors that are meant to be pulled, pull doors that should be pushed, and walk into doors that neither pull nor push, but slide. Moreover, I see others having the same troubles—unnecessary troubles. My problems with doors have become so well known that confusing doors are often called “Norman doors.” Imagine becoming famous for doors that don’t work right. I’m pretty sure that’s not what my parents planned for me. (Put “Norman doors” into your favorite search engine—be sure to include the quote marks: it makes for fascinating reading.)

How can such a simple thing as a door be so confusing? A door would seem to be about as simple a device as possible. There is not much you can do to a door: you can open it or shut it. Suppose you are in an office building, walking down a corridor. You come to a door. How does it open? Should you push or pull, on the left or the right? Maybe the door slides. If so, in which direction? I have seen doors that slide to the left, to the right, and even up into the ceiling.



FIGURE 1.1. Coffeepot for Masochists. The French artist Jacques Carelman in his series of books *Catalogue d'objets introuvables* (Catalog of unfindable objects) provides delightful examples of everyday things that are deliberately unworkable, outrageous, or otherwise ill-formed. One of my favorite items is what he calls “coffeepot for masochists.” The photograph shows a copy given to me by colleagues at the University of California, San Diego. It is one of my treasured art objects. (Photograph by Aymin Shamma for the author.)

The design of the door should indicate how to work it without any need for signs, certainly without any need for trial and error.

A friend told me of the time he got trapped in the doorway of a post office in a European city. The entrance was an imposing row of six glass swinging doors, followed immediately by a second, identical row. That’s a standard design: it helps reduce the airflow and thus maintain the indoor temperature of the building. There was no visible hardware: obviously the doors could swing in either direction: all a person had to do was push the side of the door and enter.

My friend pushed on one of the outer doors. It swung inward, and he entered the building. Then, before he could get to the next row of doors, he was distracted and turned around for an instant. He didn’t realize it at the time, but he had moved slightly to the right. So when he came to the next door and pushed it, nothing happened. “Hmm,” he thought, “must be locked.” So he pushed the side of the adjacent door. Nothing. Puzzled, my friend decided to go outside again. He turned around and pushed against the side of a door. Nothing. He pushed the adjacent door. Nothing. The door he had just entered no longer worked. He turned around once more and tried the inside doors again. Nothing. Concern, then mild panic. He was trapped! Just then, a group of people on the other side of the entranceway (to my friend’s right) passed easily through both sets of doors. My friend hurried over to follow their path.

How could such a thing happen? A swinging door has two sides. One contains the supporting pillar and the hinge, the other is unsupported. To open the door, you must push or pull on the unsupported edge. If you push on the hinge side, nothing happens. In my friend's case, he was in a building where the designer aimed for beauty, not utility. No distracting lines, no visible pillars, no visible hinges. So how can the ordinary user know which side to push on? While distracted, my friend had moved toward the (invisible) supporting pillar, so he was pushing the doors on the hinged side. No wonder nothing happened. Attractive doors. Stylish. Probably won a design prize.

Two of the most important characteristics of good design are *discoverability* and *understanding*. Discoverability: Is it possible to even figure out what actions are possible and where and how to perform them? Understanding: What does it all mean? How is the product supposed to be used? What do all the different controls and settings mean?

The doors in the story illustrate what happens when discoverability fails. Whether the device is a door or a stove, a mobile phone or a nuclear power plant, the relevant components must be visible, and they must communicate the correct message: What actions are possible? Where and how should they be done? With doors that push, the designer must provide signals that naturally indicate where to push. These need not destroy the aesthetics. Put a vertical plate on the side to be pushed. Or make the supporting pillars visible. The vertical plate and supporting pillars are natural signals, naturally interpreted, making it easy to know just what to do: no labels needed.

With complex devices, discoverability and understanding require the aid of manuals or personal instruction. We accept this if the device is indeed complex, but it should be unnecessary for simple things. Many products defy understanding simply because they have too many functions and controls. I don't think that simple home appliances—stoves, washing machines, audio and television sets—should look like Hollywood's idea of a spaceship control room. They already do, much to our consternation. Faced

with a bewildering array of controls and displays, we simply memorize one or two fixed settings to approximate what is desired.

In England I visited a home with a fancy new Italian washer-dryer combination, with super-duper multisymbol controls, all to do everything anyone could imagine doing with the washing and drying of clothes. The husband (an engineering psychologist) said he refused to go near it. The wife (a physician) said she had simply memorized one setting and tried to ignore the rest. I asked to see the manual: it was just as confusing as the device. The whole purpose of the design is lost.

The Complexity of Modern Devices

All artificial things are designed. Whether it is the layout of furniture in a room, the paths through a garden or forest, or the intricacies of an electronic device, some person or group of people had to decide upon the layout, operation, and mechanisms. Not all designed things involve physical structures. Services, lectures, rules and procedures, and the organizational structures of businesses and governments do not have physical mechanisms, but their rules of operation have to be designed, sometimes informally, sometimes precisely recorded and specified.

But even though people have designed things since prehistoric times, the field of design is relatively new, divided into many areas of specialty. Because everything is designed, the number of areas is enormous, ranging from clothes and furniture to complex control rooms and bridges. This book covers everyday things, focusing on the interplay between technology and people to ensure that the products actually fulfill human needs while being understandable and usable. In the best of cases, the products should also be delightful and enjoyable, which means that not only must the requirements of engineering, manufacturing, and ergonomics be satisfied, but attention must be paid to the entire experience, which means the aesthetics of form and the quality of interaction. The major areas of design relevant to this book are industrial design, interaction design, and experience design. None of the fields is well defined, but the focus of the efforts does vary, with industrial

designers emphasizing form and material, interactive designers emphasizing understandability and usability, and experience designers emphasizing the emotional impact. Thus:

Industrial design: The professional service of creating and developing concepts and specifications that optimize the function, value, and appearance of products and systems for the mutual benefit of both user and manufacturer (from the *Industrial Design Society of America's* website).

Interaction design: The focus is upon how people interact with technology. The goal is to enhance people's understanding of what can be done, what is happening, and what has just occurred. Interaction design draws upon principles of psychology, design, art, and emotion to ensure a positive, enjoyable experience.

Experience design: The practice of designing products, processes, services, events, and environments with a focus placed on the quality and enjoyment of the total experience.

Design is concerned with how things work, how they are controlled, and the nature of the interaction between people and technology. When done well, the results are brilliant, pleasurable products. When done badly, the products are unusable, leading to great frustration and irritation. Or they might be usable, but force us to behave the way the product wishes rather than as we wish.

Machines, after all, are conceived, designed, and constructed by people. By human standards, machines are pretty limited. They do not maintain the same kind of rich history of experiences that people have in common with one another, experiences that enable us to interact with others because of this shared understanding. Instead, machines usually follow rather simple, rigid rules of behavior. If we get the rules wrong even slightly, the machine does what it is told, no matter how insensible and illogical. People are imaginative and creative, filled with common sense; that is, a lot of valuable knowledge built up over years of experience. But instead of capitalizing on these strengths, machines require us to be precise and accurate, things we are not very good at. Machines have no

leeway or common sense. Moreover, many of the rules followed by a machine are known only by the machine and its designers.

When people fail to follow these bizarre, secret rules, and the machine does the wrong thing, its operators are blamed for not understanding the machine, for not following its rigid specifications. With everyday objects, the result is frustration. With complex devices and commercial and industrial processes, the resulting difficulties can lead to accidents, injuries, and even deaths. It is time to reverse the situation: to cast the blame upon the machines and their design. It is the machine and its design that are at fault. It is the duty of machines and those who design them to understand people. It is not our duty to understand the arbitrary, meaningless dictates of machines.

The reasons for the deficiencies in human-machine interaction are numerous. Some come from the limitations of today's technology. Some come from self-imposed restrictions by the designers, often to hold down cost. But most of the problems come from a complete lack of understanding of the design principles necessary for effective human-machine interaction. Why this deficiency? Because much of the design is done by engineers who are experts in technology but limited in their understanding of people. "We are people ourselves," they think, "so we understand people." But in fact, we humans are amazingly complex. Those who have not studied human behavior often think it is pretty simple. Engineers, moreover, make the mistake of thinking that logical explanation is sufficient: "If only people would read the instructions," they say, "everything would be all right."

Engineers are trained to think logically. As a result, they come to believe that all people must think this way, and they design their machines accordingly. When people have trouble, the engineers are upset, but often for the wrong reason. "What are these people doing?" they will wonder. "Why are they doing that?" The problem with the designs of most engineers is that they are too logical. We have to accept human behavior the way it is, not the way we would wish it to be.

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.

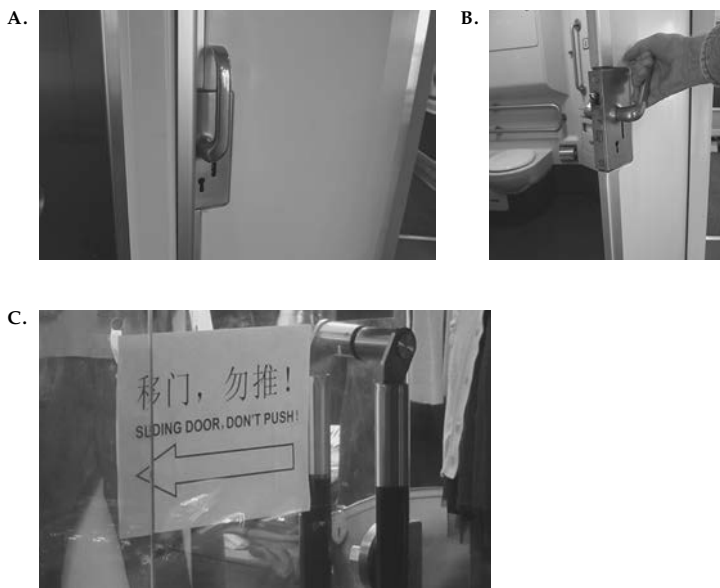


FIGURE 1.3. Sliding Doors: Seldom Done Well. Sliding doors are seldom signified properly. The top two photographs show the sliding door to the toilet on an Amtrak train in the United States. The handle clearly signifies "pull," but in fact, it needs to be rotated and the door slid to the right. The owner 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.

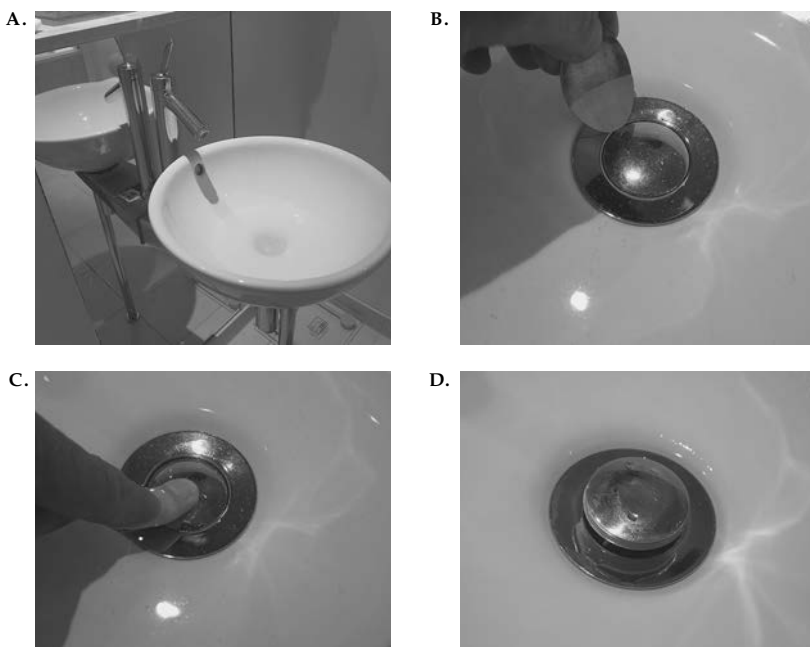
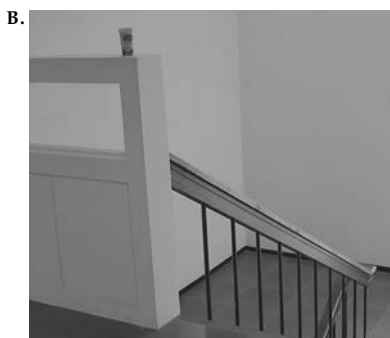


FIGURE 1.4. The Sink That Would Not Drain: Where Signifiers Fail. I washed my hands in my hotel sink in London, but then, as shown in Photo A, was left with the question of how to empty the sink of the dirty water. I searched all over for a control: none. I tried prying open the sink stopper with a spoon (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

the device, for that is where it is displayed and manipulated. But in fact, in many cases the actual material is “in the cloud,” located on some distant machine. The conceptual model is of one, coherent image, whereas it may actually consist of parts, each located on different machines that could be almost anywhere in the world. This simplified model is helpful for normal usage, but if the network connection to the cloud services is interrupted, the result can be confusing. Information is still on their screen, but users can no longer save it or retrieve new things: their conceptual model offers no explanation. Simplified models are valuable only as long as the assumptions that support them hold true.

There are often multiple conceptual models of a product or device. People’s conceptual models for the way that regenerative braking in a hybrid or electrically powered automobile works are quite different for average drivers than for technically sophisticated drivers, different again for whoever must service the system, and yet different again for those who designed the system.

Conceptual models found in technical manuals and books for technical use can be detailed and complex. The ones we are concerned with here are simpler: they reside in the minds of the people who are using the product, so they are also “mental models.” Mental models, as the name implies, are the conceptual models in people’s minds that represent their understanding of how things work. Different people may hold different mental models of the same item. Indeed, a single person might have multiple models of the same item, each dealing with a different aspect of its operation: the models can even be in conflict.

Conceptual models are often inferred from the device itself. Some models are passed on from person to person. Some come from manuals. Usually the device itself offers very little assistance, so the model is constructed by experience. Quite often these models are erroneous, and therefore lead to difficulties in using the device.

The major clues to how things work come from their perceived structure—in particular from signifiers, affordances, constraints, and mappings. Hand tools for the shop, gardening, and the house tend to make their critical parts sufficiently visible that concep-



FIGURE 1.8. Junghans Mega 1000 Digital Radio Controlled Watch. There is no good conceptual model for understanding the operation of my watch. It has five buttons with no hints as to what each one does. And yes, the buttons do different things in their different modes. But it is a very nice-looking watch, and always has the exact time because it checks official radio time stations. (The top row of the display is the date: Wednesday, February 20, the eighth week of the year.) (Photograph by the author.)

tual models of their operation and function are readily derived. Consider a pair of scissors: you can see that the number of possible actions is limited. The holes are clearly there to put something into, and the only logical things that will fit are fingers. The holes are both affordances—they allow the fingers to be inserted—and signifiers—they indicate where the fingers are to go. The sizes of the holes provide constraints to limit the possible fingers: a big hole suggests several fingers; a small hole, only one. The mapping between holes and fingers—the set of possible operations—is signified and constrained by the holes. Moreover, the operation is not sensitive to finger placement: if you use the wrong fingers (or the wrong hand), the scissors still work, although not as comfortably. You can figure out the scissors because their operating parts are visible and the implications clear. The conceptual model is obvious, and there is effective use of signifiers, affordances, and constraints.

What happens when the device does not suggest a good conceptual model? Consider my digital watch with five buttons: two along the top, two along the bottom, and one on the left side (Figure 1.8). What is each button for? How would you set the time? There is no way to tell—no evident relationship between the operating controls and the functions, no constraints, no apparent mappings. Moreover, the buttons have multiple ways of being used. Two of the buttons do different things when pushed quickly or when kept depressed for several seconds. Some operations require simultaneous depression of several of the buttons. The only way to tell how to work the watch is to read the manual, over and over again. With the scissors, moving the handle makes the blades move. The watch provides no

visible relationship between the buttons and the possible actions, no discernible relationship between the actions and the end results. I really like the watch: too bad I can't remember all the functions.

Conceptual models are valuable in providing understanding, in predicting how things will behave, and in figuring out what to do when things do not go as planned. A good conceptual model allows us to predict the effects of our actions. Without a good model, we operate by rote, blindly; we do operations as we were told to do them; we can't fully appreciate why, what effects to expect, or what to do if things go wrong. As long as things work properly, we can manage. When things go wrong, however, or when we come upon a novel situation, then we need a deeper understanding, a good model.

For everyday things, conceptual models need not be very complex. After all, scissors, pens, and light switches are pretty simple devices. There is no need to understand the underlying physics or chemistry of each device we own, just the relationship between the controls and the outcomes. When the model presented to us is inadequate or wrong (or, worse, nonexistent), we can have difficulties. Let me tell you about my refrigerator.

I used to own an ordinary, two-compartment refrigerator—nothing very fancy about it. The problem was that I couldn't set the temperature properly. There were only two things to do: adjust the temperature of the freezer compartment and adjust the tempera-

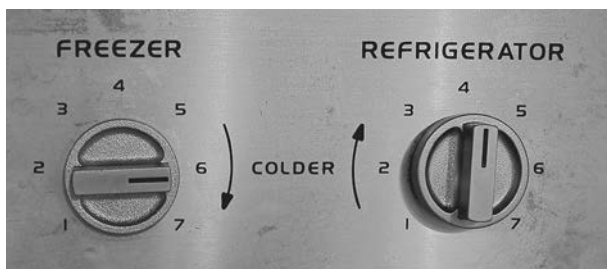


FIGURE 1.9. Refrigerator Controls. Two compartments—fresh food and freezer—and two controls (in the fresh food section just right). Your task: Suppose the freezer is too cold, the fresh food section just right. How would you adjust the controls so as to make the freezer warmer and keep the fresh food the same? (Photograph by the author.)

ture of the fresh food compartment. And there were two controls, one labeled “freezer,” the other “refrigerator.” What’s the problem?

Oh, perhaps I’d better warn you. The two controls are not independent. The freezer control also affects the fresh food temperature, and the fresh food control also affects the freezer. Moreover, the manual warns that one should “always allow twenty-four (24) hours for the temperature to stabilize whether setting the controls for the first time or making an adjustment.”

It was extremely difficult to regulate the temperature of my old refrigerator. Why? Because the controls suggest a false conceptual model. Two compartments, two controls, which implies that each control is responsible for the temperature of the compartment that carries its name: this conceptual model is shown in Figure 1.10A. It is wrong. In fact, there is only one thermostat and only one cooling mechanism. One control adjusts the thermostat setting, the other the relative proportion of cold air sent to each of the two compartments of the refrigerator. This is why the two controls interact: this conceptual model is shown in Figure 1.10B. In addition, there must be a temperature sensor, but there is no way of knowing where it is located. With the conceptual model suggested by the controls,

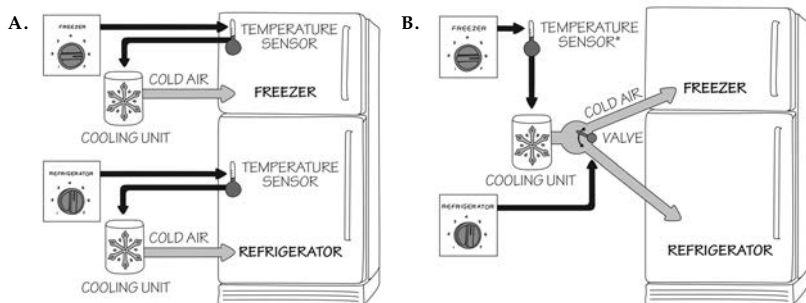


FIGURE 1.10. Two Conceptual Models for a Refrigerator. The conceptual model A is provided by the system image of the refrigerator as gleaned from the controls. Each control determines the temperature of the named part of the refrigerator. This means that each compartment has its own temperature sensor and cooling unit. This is wrong. The correct conceptual model is shown in B. There is no way of knowing where the temperature sensor is located so it is shown outside the refrigerator. The freezer control determines the freezer temperature (so is this where the sensor is located?). The refrigerator control determines how much of the cold air goes to the freezer and how much to the refrigerator.

adjusting the temperatures is almost impossible and always frustrating. Given the correct model, life would be much easier.

Why did the manufacturer suggest the wrong conceptual model? We will never know. In the twenty-five years since the publication of the first edition of this book, I have had many letters from people thanking me for explaining their confusing refrigerator, but never any communication from the manufacturer (General Electric). Perhaps the designers thought the correct model was too complex, that the model they were giving was easier to understand. But with the wrong conceptual model, it was impossible to set the controls. And even though I am convinced I knew the correct model, I still couldn't accurately adjust the temperatures because the refrigerator design made it impossible to discover which control was for the temperature sensor, which for the relative proportion of cold air, and in which compartment the sensor was located. The lack of immediate feedback for the actions did not help: it took twenty-four hours to see whether the new setting was appropriate. I shouldn't have to keep a laboratory notebook and do controlled experiments just to set the temperature of my refrigerator.

I am happy to say that I no longer own that refrigerator. Instead I have one that has two separate controls, one in the fresh food compartment, one in the freezer compartment. Each control is nicely calibrated in degrees and labeled with the name of the compartment it controls. The two compartments are independent: setting the temperature in one has no effect on the temperature in the other. This solution, although ideal, does cost more. But far less expensive solutions are possible. With today's inexpensive sensors and motors, it should be possible to have a single cooling unit with a motor-controlled valve controlling the relative proportion of cold air diverted to each compartment. A simple, inexpensive computer chip could regulate the cooling unit and valve position so that the temperatures in the two compartments match their targets. A bit more work for the engineering design team? Yes, but the results would be worth it. Alas, General Electric is still selling refrigerators with the very same controls and mechanisms that cause so much

confusion. The photograph in Figure 1.9 is from a contemporary refrigerator, photographed in a store while preparing this book.

The System Image

People create mental models of themselves, others, the environment, and the things with which they interact. These are conceptual models formed through experience, training, and instruction. These models serve as guides to help achieve our goals and in understanding the world.

How do we form an appropriate conceptual model for the devices we interact with? We cannot talk to the designer, so we rely upon whatever information is available to us: what the device looks like, what we know from using similar things in the past, what was told to us in the sales literature, by salespeople and advertisements, by articles we may have read, by the product website and instruction manuals. I call the combined information available to us the *system image*. When the system image is incoherent or inappropriate, as in the case of the refrigerator, then the user cannot easily use the device. If it is incomplete or contradictory, there will be trouble.

As illustrated in Figure 1.11, the designer of the product and the person using the product form somewhat disconnected vertices of a triangle. The designer's conceptual model is the designer's conception of the product, occupying one vertex of the triangle. The product itself is no longer with the designer, so it is isolated as a second vertex, perhaps sitting on the user's kitchen counter. The system image is what can be perceived from the physical structure that has been built (including documentation, instructions, signifiers, and any information available from websites and help lines). The user's conceptual model comes from the system image, through interaction with the product, reading, searching for online information, and from whatever manuals are provided. The designer expects the user's model to be identical to the design model, but because designers cannot communicate directly with users, the entire burden of communication is on the system image.

FIGURE 1.11. The Designer's Model, the User's Model, and the System Image. The designer's conceptual model is the designer's conception of the look, feel, and operation of a product. The system image is what can be derived from the physical structure that has been built (including documentation). The user's mental model is developed through interaction with the product and the system image. Designers expect the user's model to be identical to their own, but because they cannot communicate directly with the user, the burden of communication is with the system image.



Figure 1.11 indicates why communication is such an important aspect of good design. No matter how brilliant the product, if people cannot use it, it will receive poor reviews. It is up to the designer to provide the appropriate information to make the product understandable and usable. Most important is the provision of a good conceptual model that guides the user when things go wrong. With a good conceptual model, people can figure out what has happened and correct the things that went wrong. Without a good model, they struggle, often making matters worse.

Good conceptual models are the key to understandable, enjoyable products: good communication is the key to good conceptual models.

The Paradox of Technology

Technology offers the potential to make life easier and more enjoyable; each new technology provides increased benefits. At the same time, added complexities increase our difficulty and frustration with technology. The design problem posed by technological advances is enormous. Consider the wristwatch. A few decades ago, watches were simple. All you had to do was set the time and keep the watch wound. The standard control was the stem: a knob at the side of the watch. Turning the knob would wind the spring that provided power to the watch movement. Pulling out the knob and turning it rotated the hands. The operations were easy to learn and easy to do. There was a reasonable relationship between the

turning of the knob and the resulting turning of the hands. The design even took into account human error. In its normal position, turning the stem wound the mainspring of the clock. The stem had to be pulled before it would engage the gears for setting the time. Accidental turns of the stem did no harm.

Watches in olden times were expensive instruments, manufactured by hand. They were sold in jewelry stores. Over time, with the introduction of digital technology, the cost of watches decreased rapidly, while their accuracy and reliability increased. Watches became tools, available in a wide variety of styles and shapes and with an ever-increasing number of functions. Watches were sold everywhere, from local shops to sporting goods stores to electronic stores. Moreover, accurate clocks were incorporated in many appliances, from phones to musical keyboards: many people no longer felt the need to wear a watch. Watches became inexpensive enough that the average person could own multiple watches. They became fashion accessories, where one changed the watch with each change in activity and each change of clothes.

In the modern digital watch, instead of winding the spring, we change the battery, or in the case of a solar-powered watch, ensure that it gets its weekly dose of light. The technology has allowed more functions: the watch can give the day of the week, the month, and the year; it can act as a stopwatch (which itself has several functions), a countdown timer, and an alarm clock (or two); it has the ability to show the time for different time zones; it can act as a counter and even as a calculator. My watch, shown in Figure 1.8, has many functions. It even has a radio receiver to allow it to set its time with official time stations around the world. Even so, it is far less complex than many that are available. Some watches have built-in compasses and barometers, accelerometers, and temperature gauges. Some have GPS and Internet receivers so they can display the weather and news, e-mail messages, and the latest from social networks. Some have built-in cameras. Some work with buttons, knobs, motion, or speech. Some detect gestures. The watch is no longer just an instrument for telling time: it has become a platform for enhancing multiple activities and lifestyles.

The added functions cause problems: How can all these functions fit into a small, wearable size? There are no easy answers. Many people have solved the problem by not using a watch. They use their phone instead. A cell phone performs all the functions much better than the tiny watch, while also displaying the time.

Now imagine a future where instead of the phone replacing the watch, the two will merge, perhaps worn on the wrist, perhaps on the head like glasses, complete with display screen. The phone, watch, and components of a computer will all form one unit. We will have flexible displays that show only a tiny amount of information in their normal state, but that can unroll to considerable size. Projectors will be so small and light that they can be built into watches or phones (or perhaps rings and other jewelry), projecting their images onto any convenient surface. Or perhaps our devices won't have displays, but will quietly whisper the results into our ears, or simply use whatever display happens to be available: the display in the seatback of cars or airplanes, hotel room televisions, whatever is nearby. The devices will be able to do many useful things, but I fear they will also frustrate: so many things to control, so little space for controls or signifiers. The obvious solution is to use exotic gestures or spoken commands, but how will we learn, and then remember, them? As I discuss later, the best solution is for there to be agreed upon standards, so we need learn the controls only once. But as I also discuss, agreeing upon these is a complex process, with many competing forces hindering rapid resolution. We will see.

The same technology that simplifies life by providing more functions in each device also complicates life by making the device harder to learn, harder to use. This is the paradox of technology and the challenge for the designer.

The Design Challenge

Design requires the cooperative efforts of multiple disciplines. The number of different disciplines required to produce a successful product is staggering. Great design requires great designers, but that isn't enough: it also requires great management, because the

hardest part of producing a product is coordinating all the many, separate disciplines, each with different goals and priorities. Each discipline has a different perspective of the relative importance of the many factors that make up a product. One discipline argues that it must be usable and understandable, another that it must be attractive, yet another that it has to be affordable. Moreover, the device has to be reliable, be able to be manufactured and serviced. It must be distinguishable from competing products and superior in critical dimensions such as price, reliability, appearance, and the functions it provides. Finally, people have to actually purchase it. It doesn't matter how good a product is if, in the end, nobody uses it.

Quite often each discipline believes its distinct contribution to be most important: "Price," argues the marketing representative, "price plus these features." "Reliable," insist the engineers. "We have to be able to manufacture it in our existing plants," say the manufacturing representatives. "We keep getting service calls," say the support people; "we need to solve those problems in the design." "You can't put all that together and still have a reasonable product," says the design team. Who is right? Everyone is right. The successful product has to satisfy all these requirements.

The hard part is to convince people to understand the viewpoints of the others, to abandon their disciplinary viewpoint and to think of the design from the viewpoints of the person who buys the product and those who use it, often different people. The viewpoint of the business is also important, because it does not matter how wonderful the product is if not enough people buy it. If a product does not sell, the company must often stop producing it, even if it is a great product. Few companies can sustain the huge cost of keeping an unprofitable product alive long enough for its sales to reach profitability—with new products, this period is usually measured in years, and sometimes, as with the adoption of high-definition television, decades.

Designing well is not easy. The manufacturer wants something that can be produced economically. The store wants something that will be attractive to its customers. The purchaser has several

demands. In the store, the purchaser focuses on price and appearance, and perhaps on prestige value. At home, the same person will pay more attention to functionality and usability. The repair service cares about maintainability: how easy is the device to take apart, diagnose, and service? The needs of those concerned are different and often conflict. Nonetheless, if the design team has representatives from all the constituencies present at the same time, it is often possible to reach satisfactory solutions for all the needs. It is when the disciplines operate independently of one another that major clashes and deficiencies occur. The challenge is to use the principles of human-centered design to produce positive results, products that enhance lives and add to our pleasure and enjoyment. The goal is to produce a great product, one that is successful, and that customers love. It can be done.